

## THE PHOTOECCENTRIC EFFECT AND PROTO-HOT JUPITERS. III. A PAUCITY OF PROTO-HOT JUPITERS ON SUPER-ECCENTRIC ORBITS

REBEKAH I. DAWSON<sup>1,2,6</sup>, RUTH A. MURRAY-CLAY<sup>2,3</sup>, AND JOHN ASHER JOHNSON<sup>2,4,5</sup>

<sup>1</sup> Department of Astronomy, University of California, Berkeley, Hearst Field Annex B-20, Berkeley, CA 94720-3411, USA

<sup>2</sup> Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, MS-10, Cambridge, MA 02138, USA; [rdawson@berkeley.edu](mailto:rdawson@berkeley.edu)

<sup>3</sup> Department of Physics, University of California, Santa Barbara, Broida Hall, Santa Barbara, CA 93106-9530, USA

<sup>4</sup> Division of Geological and Planetary Sciences, California Institute of Technology, 1200 East California Boulevard, MC 170-25, Pasadena, CA 91125, USA

<sup>5</sup> NASA Exoplanet Science Institute (NExScI), CIT Mail Code 100-22, 770 South Wilson Avenue, Pasadena, CA 91125, USA

Received 2012 October 26; accepted 2014 October 20; published 2014 December 22

### ABSTRACT

Gas giant planets orbiting within 0.1 AU of their host stars are unlikely to have formed in situ and are evidence for planetary migration. It is debated whether the typical hot Jupiter smoothly migrated inward from its formation location through the proto-planetary disk, or was perturbed by another body onto a highly eccentric orbit, which tidal dissipation subsequently shrank and circularized during close stellar passages. Socrates and collaborators predicted that the latter model should produce a population of super-eccentric proto-hot Jupiters readily observable by *Kepler*. We find a paucity of such planets in the *Kepler* sample, which is inconsistent with the theoretical prediction with 96.9% confidence. Observational effects are unlikely to explain this discrepancy. We find that the fraction of hot Jupiters with an orbital period  $P > 3$  days produced by the star–planet Kozai mechanism does not exceed (at two-sigma) 44%. Our results may indicate that disk migration is the dominant channel for producing hot Jupiters with  $P > 3$  days. Alternatively, the typical hot Jupiter may have been perturbed to a high eccentricity by interactions with a planetary rather than stellar companion, and began tidal circularization much interior to 1 AU after multiple scatterings. A final alternative is that early in the tidal circularization process at high eccentricities tidal circularization occurs much more rapidly than later in the process at low eccentricities, although this is contrary to current tidal theories.

**Key word:** planetary systems

### 1. INTRODUCTION

Roughly 1% of Sun-like stars host hot Jupiters, giant planets with small semi-major axes (Mayor et al. 2011; Howard et al. 2012; Wright et al. 2012). Unlikely to have formed in situ (Rafikov 2006), hot Jupiters are evidence for the prevalence of planetary migration, which may take place via interactions with the proto-planetary disk (e.g., Goldreich & Tremaine 1980; Ward 1997; Alibert et al. 2005; Ida & Lin 2008; Bromley & Kenyon 2011), or other bodies in the system. One or more companions can create a hot Jupiter by perturbing a cold Jupiter onto an eccentric orbit, which tidal forces then shrink and circularize during close passages to the star. Proposed mechanisms for this high eccentricity migration (HEM) include Kozai oscillations induced by a distant stellar binary companion (star–planet Kozai; e.g., Wu & Murray 2003; Fabrycky & Tremaine 2007; Naoz et al. 2012) or by another planet in the system (planet–planet Kozai; Naoz et al. 2011; Lithwick & Naoz 2011), planet–planet scattering (e.g., Rasio & Ford 1996; Ford & Rasio 2006, 2008; Chatterjee et al. 2008; Jurić & Tremaine 2008; Matsumura et al. 2010; Nagasawa & Ida 2011; Beaugé & Nesvorný 2012; Boley et al. 2012), and secular chaos (Wu & Lithwick 2011). We consider interactions with other bodies in the system to also encompass gravitational perturbations preceded by disk migration (e.g., Guillochon et al. 2011).

One way to distinguish whether disk migration or HEM is dominant in setting the architecture of systems of giant planets is to search for other populations of giant planets, in addition to hot Jupiters, that may also result from HEM, including (1) failed

hot Jupiters, which are stranded at high eccentricities but with periapses too large to undergo significant tidal circularization over the star’s lifetime, (2) Jupiters on short-period, moderately eccentric orbits, nearing the end of their HEM journey, and (3) proto-hot Jupiters on super-eccentric orbits in the process of HEM. Recently, Socrates et al. (2012b, S12 hereafter) suggested that if HEM is the dominant channel for producing hot Jupiters, we should readily detect a number of super-eccentric Jupiters in the act of migrating inward. Moreover, they showed that the number of super-eccentric Jupiters can be estimated from the number of moderately eccentric Jupiters that have similar angular momentum, based on their relative circularization rates. Based on the number of moderately eccentric, short-period Jupiters found by other planet hunting programs (tabulated in the Exoplanet Orbit Database, EOD, by Wright et al. 2011), S12 predicted that the *Kepler* Mission should discover five to seven proto-hot Jupiters with eccentricities  $e > 0.9$  and noted that these planets should in fact already be present in the Borucki et al. (2011) candidate collection.

The S12 prediction requires the steady production rate of hot Jupiters throughout the Galaxy to be represented in the observed sample, as well as several conventional assumptions including the conservation of the migrating Jupiter’s angular momentum, tidal circularization under the constant time lag approximation, and the beginning of HEM at or beyond an orbital period of 2 yr. This prediction is a useful, quantitative test for discerning the origin of hot Jupiters. Confirmation of the S12 prediction would reveal that hot Jupiters are placed on their close-in orbits by interactions with companions, not a disk, while a paucity of proto-hot Jupiters in the *Kepler* sample would inform us that HEM is not the dominant channel, or that some aspect of our current understanding of HEM is incorrect.

<sup>6</sup> Miller Fellow.

Motivated by the S12 prediction, we use what we term the photoeccentric effect to measure the eccentricities of Jupiter-sized planets from their transit light curves (Dawson & Johnson 2012; hereafter DJ12). We validated our approach by comparing our eccentricity measurement obtained from the light curve to (when available) radial-velocity measurements, including for HD-17156-b (DJ12), Kepler-419-b (formerly KOI-1474.01, Dawson et al. 2014), KOI-868 B (DJ12; Díaz et al. 2014), and KOI-889-b.<sup>7</sup> Prior to the radial velocity follow-up, Kepler-419-b was identified by Dawson et al. (2012, hereafter D12) as a transiting planet candidate with a long orbital period (69.7 days), a large eccentricity ( $e = 0.81 \pm 0.10$ ), and transit timing variations caused by a massive outer companion. Originally, uncertainty in the candidate's eccentricity made it ambiguous whether Kepler-419-b was one of the proto-hot Jupiters predicted by S12 or, alternatively, a failed-hot Jupiter beyond the reach of tidal circularization over its host star's lifetime. It was later shown to be failed-hot Jupiter (Dawson et al. 2014).

Here we examine the entire sample of *Kepler* Jupiters to assess whether the planets expected from HEM are present. We find with 96.9% confidence that the putative highly eccentric progenitors of hot Jupiters are partly or entirely missing from the *Kepler* sample. In Section 2, we update the S12 prediction, accounting for Poisson counting uncertainties and incompleteness, and translate it into a prediction for transit light curve observables. In Section 3, we compare the prediction of Section 2 to the light curve properties of candidates in the *Kepler* sample, and conclude that there is a paucity of proto-hot Jupiters. In Section 4, we place an upper limit on the fraction of hot Jupiters created by stellar binaries, consider the contribution of disk migration to the hot Jupiter population, and present Monte Carlo predictions for other dynamical scenarios, finding that the paucity of proto-hot Jupiters can be compatible with HEM. We conclude (Section 5) by outlining the theoretical and observational pathways necessary to distinguish the dominant channel for hot Jupiter creation.

## 2. UPDATED PREDICTION FOR NUMBER OF SUPER-ECCENTRIC PROTO-HOT JUPITERS AND TRANSIT LIGHT CURVE OBSERVABLES

In Section 2.1, we derive the expected number of identifiable *Kepler* super-eccentric proto-hot Jupiters, following S12 but using updated survey samples. We refine the S12 prediction by quantifying its uncertainty and incorporating incompleteness. In Section 2.2, we describe how to confirm or rule out the existence of super-eccentric proto-hot Jupiters using *Kepler* photometry alone, by recasting the prediction in terms of light curve observables.

### 2.1. Expected Number of Proto-hot Jupiters with $e > 0.9$ in the *Kepler* Sample

S12 predicted that the *Kepler* mission should discover a number of super-eccentric hot Jupiter progenitors in the process of HEM. Previously (DJ12), we showed that super-eccentric planets should be easily identifiable from their transit light curves, and thus precise radial-velocity (RV) follow-up is unnecessary. This is fortunate because most *Kepler* stars are too faint to be amenable to precise RV observations. To predict the

number of super-eccentric Jupiters, S12 considered a population of proto-hot Jupiters undergoing tidal circularization along a track of constant angular momentum defined by a final semi-major axis,  $a_{\text{final}} = a(1 - e^2)$ , and related by Kepler's law, a final orbital period  $P_{\text{final}} = P(1 - e^2)^{3/2}$  for which  $a$ ,  $e$ , and  $P$  are the observed semi-major axis, eccentricity, and orbital period, respectively. The number of super-eccentric Jupiters ( $e > 0.9$ ,  $\bar{N}_{\text{sup}}$ ) along a track of constant angular momentum is related to the number of moderately eccentric Jupiters ( $0.2 < e < 0.6$ ,  $\bar{N}_{\text{mod}}$ ) along the same track by:

$$\bar{N}_{\text{sup}} = \bar{N}_{\text{mod}} r(e_{\text{max}}), \quad (1)$$

where the variable  $e_{\text{max}} = [1 - (P_{\text{final}}/P_{\text{max}})^{2/3}]^{1/2}$  is set by the maximum observable orbital period  $P_{\text{max}}$  and  $r(e_{\text{max}})$  is the ratio of time spent at super eccentricities ( $0.9 < e < e_{\text{max}}$ ) to moderate eccentricities ( $0.2 < e < 0.6$ ). Bars over  $N_{\text{sup}}$  and  $N_{\text{mod}}$  indicate that these are mean numbers. The observationally counted numbers are sampled from Poisson distributions defined by these means.

Most Jupiters in the *Kepler* sample lack measured eccentricities, so the  $\bar{N}_{\text{mod}}$  of the *Kepler* sample is unknown. Following S12, we use the sample of planets detected by non-*Kepler* surveys, which we denote with subscript 0 (Figure 1). To estimate  $\bar{N}_{\text{mod}}$  along a track in the *Kepler* sample, we use the ratio of  $\bar{N}_{\text{mod},0}$  to the number in another class of calibration object. This other class must be countable in the *Kepler* sample, and would ideally be along a  $P_{\text{final}}$  track. However, because the eccentricities of the *Kepler* planets are unknown, we use the planets with an orbital period  $P = P_{\text{final}}$ , of which there are  $\bar{N}_{P=P_{\text{final}},0}$  in the calibration sample. If we assume the ratio  $\bar{N}_{\text{mod},0}/\bar{N}_{P=P_{\text{final}},0}$  is the same for the calibration sample as for the *Kepler* sample, then we can compute the expected  $\bar{N}_{\text{mod}}$  for the *Kepler* sample:

$$\bar{N}_{\text{mod}} = \frac{\bar{N}_{\text{mod},0}}{\bar{N}_{P=P_{\text{final}},0}} \bar{N}_{P=P_{\text{final}}}. \quad (2)$$

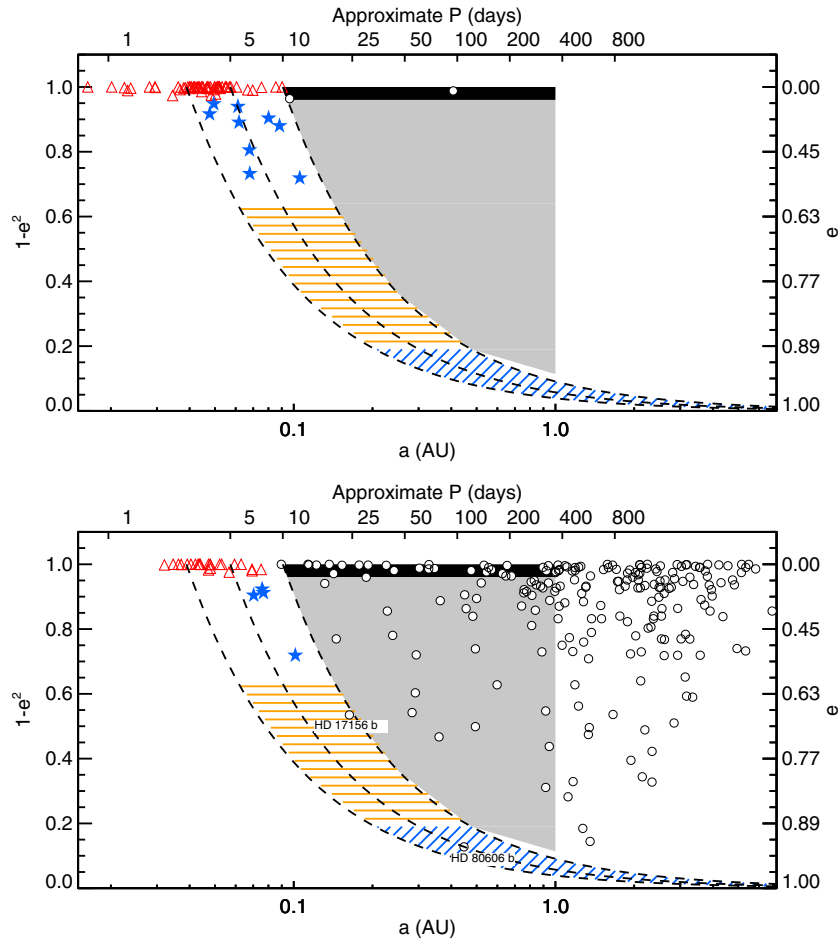
The prediction by S12 was for an ideal *Kepler* sample complete out to orbital periods of 2 yr (S. Dong 2012, private communication). The *Kepler* mission was originally designed to operate for 3 yr, and operated for 4 yr before the reaction wheels failed. To derive the expected number of super-eccentric proto-hot Jupiters in a sample of a limited timespan  $t_{\text{survey}}$ , we must account for incompleteness. If  $N_{\text{trans,min}}$  transits are the minimum number of transits required for the *Kepler* transit pipeline to detect the proto-hot Jupiter, the completeness (with respect to this effect alone)  $C_{\text{comp}}$  ranges from 100% at orbital periods  $\leq t_{\text{survey}}/N_{\text{trans,min}}$  to 0% at orbital periods of  $t_{\text{survey}}/(N_{\text{trans,min}} - 1)$ . We define  $r(e_{\text{max}})$  to account for incompleteness. (Note that the completeness is a function of eccentricity because of the one-to-one relation between eccentricity and orbital period along a given angular momentum track.):

$$r(e_{\text{max}}) = \frac{\int_{0.9}^{e_{\text{max}}} C_{\text{comp}}(e) |\dot{e}|^{-1} de}{\int_{0.2}^{0.6} |\dot{e}|^{-1} de}, \quad (3)$$

for which  $\dot{e}$  is due to tides raised on the planet under the constant tidal time lag approximation (Eggleton et al. 1998; Hansen 2010, S12; Socrates & Katz 2012; Socrates et al. 2012a), and the completeness  $C_{\text{comp}}(e)$  is

$$C_{\text{comp}}(e) = \begin{cases} 1, & e < e_{\text{complete}} \\ 1 + (1 - e^2)^{3/2} (t_{\text{survey}}/P_{\text{final}}) - N_{\text{trans,min}}, & e_{\text{complete}} < e < e_{\text{max}}, \end{cases} \quad (4)$$

<sup>7</sup> In an unpublished work, we measured a value  $e = 0.5$ . Hébrard et al. (2013) measured a value of 0.589 from SOPHIE radial-velocity measurements.



**Figure 1.** Giant planets detected by non-*Kepler* surveys from the EOD (Wright et al. 2012) using the transit technique (top) or RV technique (bottom). The top  $x$ -axis, for reference, indicates the equivalent orbital period for a planet orbiting a solar-mass star. All are Jupiter-mass ( $M_p \sin i > 0.25 M_{\text{Jup}}$  or  $8 R_{\oplus} < R_p < 22 R_{\oplus}$ ) planets orbiting stars with  $4500 \text{ K} < T_{\text{eff}} < 6500 \text{ K}$ , and  $\log g$  within the uncertainties consistent with  $> 4$ . The dashed lines represent tracks of  $P_{\text{final}} = 2.8, 5, 10$  days. The shaded and patterned regions correspond to Figure 6. Within the 3 days  $< P_{\text{final}} < 10$  days angular momentum tracks are hot Jupiters (red triangles), moderately eccentric Jupiters with  $0.2 < e < 0.6$  (blue stars), Jupiters with  $0.6 < e < 0.9$  (orange horizontal striped region), and super-eccentric Jupiters (blue, diagonal-striped region). The RV-discovered planet HD 17156 b lies in the orange, horizontal striped region, and the RV-discovered planet HD 80606 b lies in the blue, diagonal-striped region. Period valley: the gray region denotes Jupiters with  $P_{\text{final}} > 10$  days, but interior to the uptick in giant planet frequency at 1 AU, and the black region contains circular Jupiters in this region.

and  $e_{\text{complete}} = [1 - (N_{\text{trans,min}} P_{\text{final}}/t_{\text{survey}})^{2/3}]^{1/2}$ , and  $e_{\text{max}} = (1 - [(N_{\text{trans,min}} - 1) P_{\text{final}}/t_{\text{survey}}]^{2/3})^{1/2}$ .

Although calculations are often made under the assumption that the *Kepler* candidate list (Borucki et al. 2011; Batalha et al. 2013a; Burke et al. 2014) is complete for Jupiter-sized planets exhibiting two transits in the timespan under consideration (e.g., Fressin et al. 2013), in this paper we estimate the completeness more conservatively. The *Kepler* pipeline is set up to detect objects that transit three times during the quarters over which the pipeline was run; all reported candidates that transit only one to two times were detected by eye (C. Burke & J. Rowe 2013, private communication). There is no estimate available for the completeness of detections made by eye. We compile a sample of candidates that is complete for giant planets that transit three times in Q1–Q16, cross-checking among several sources. We describe our sample of long-period *Kepler* giant planets in Appendix A. We employ Equation (4) using  $t_{\text{survey}} = 4 \text{ yr}$  and  $N_{\text{trans,min}} = 3$ . We obtain  $r = 1.05, 0.809, 0.539$  for  $P_{\text{final}} = 3, 5, 10$  days, respectively. Later in this section we will update the completeness further to account for noise and missing data.

Next we describe the selection cuts we make to count  $N_{\text{mod},0}$  (blue stars, Figure 1),  $N_{P=P_{\text{final},0}}$  (open symbols, Figure 1), and  $N_{P=P_{\text{final}}}$ . Because the stellar parameters from the *Kepler* Input Catalog (KIC) are not reliable for stars outside the temperature range  $4500 < T_{\text{eff}} < 6500 \text{ K}$  (Brown et al. 2011), we only include stars within this temperature range in both the *Kepler* and calibration samples.<sup>8</sup> We impose a cut of stellar surface gravity  $\log g > 4$  to exclude giant stars because their KIC parameters are unreliable (we include stars with  $\log g < 4$  but that are consistent with  $\log g = 4$  within two sigma). All stellar parameters are taken from Huber et al. (2014).

We select planets with  $8 R_{\oplus} < R_p < 22 R_{\oplus}$ , where the radius is calculated using  $R_p/R_*$  from the Q1–Q12 candidate list and  $R_*$  from the updated stellar parameters in Huber et al. (2014). We follow S12 and consider two  $P_{\text{final}}$  intervals: 2.8 days  $< P_{\text{final}} < 5$  days (Interval 1)<sup>9</sup> and 5 days  $< P_{\text{final}} < 10$  days (Interval 2). The transit probability does not change much throughout each

<sup>8</sup> However, RV surveys are only complete out to 6100 K, so the RV sample spans a more limited stellar temperature range.

<sup>9</sup> We use a lower limit of 2.8 days because that is the  $P_{\text{final}}$  below which we do not see any moderately eccentric Jupiters in the non-*Kepler* surveys.

**Table 1**  
Counted Planets

$e$	Interval (days)	Counted	Mean <sup>a</sup>	Sample <sup>b</sup>
0.2 < $e$ < 0.6	1: 2.8–5	$N_{\text{mod},0} = 6$	$\bar{N}_{\text{mod},0} = 6^{+3}_{-2}$	Cal
	2: 5–10	$N_{\text{mod},0} = 7$	$\bar{N}_{\text{mod},0} = 7^{+3}_{-2}$	Cal
unspecified	1: 2.8–5	$N_{P=P_{\text{final}},0} = 69$	$\bar{N}_{P=P_{\text{final}},0} = 69^{+9}_{-8}$	Cal
	1: 2.8–5	$N_{P=P_{\text{final}}} = 24$	$\bar{N}_{P=P_{\text{final}}} = 24 \pm 5$	Kep
	2: 5–10	$N_{P=P_{\text{final}},0} = 18$	$\bar{N}_{P=P_{\text{final}},0} = 18^{+5}_{-4}$	Cal
	2: 5–10	$N_{P=P_{\text{final}}} = 16$	$\bar{N}_{P=P_{\text{final}}} = 16 \pm 4$	Kep

**Notes.**

<sup>a</sup> Median, with 68.3% confidence interval, of the posterior of Poisson means, each defining a Poisson distribution from which the counted number may be sampled.

<sup>b</sup> Kep, *Kepler*; Cal, calibration non-*Kepler* (Figure 1).

interval. We tabulate the counted numbers and their sources in Table 1.

Each number of counted planets (Table 1) is drawn from a Poisson distribution with an unknown mean. We wish to compute the expected number of super-eccentric proto-hot Jupiters not using the counted numbers (which are only samples from a Poisson distribution), but rather using estimated posteriors for the mean numbers, incorporating uncertainty. We use a Jeffrey’s prior; see Appendix B for a description of our approach. Note that in the calibration sample, we exclude planets with poorly constrained eccentricities. For planets with  $e = 0$  in the EOD, we refer to the literature or fit the data ourselves, and only include planets listed with  $e = 0$  that are constrained to have  $e < 0.2$ . See Appendix C for more details.

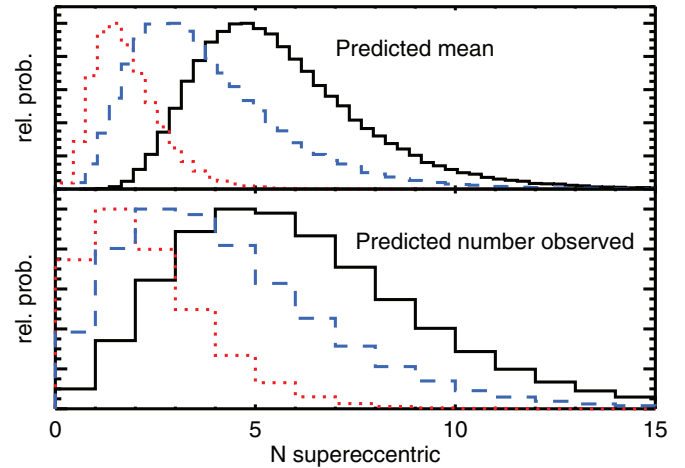
There are two additional effects on the completeness that we now consider. First, transits may fall during gaps in the data or missing quarters. To incorporate this effect, we numerically integrate Equation (3), inserting an extra factor  $C_{\text{comp,sampled}}$  into the integrand, where  $C_{\text{comp,sampled}}$  is the fraction of phases for which we would observe three or more transits during Q1–Q16.

We estimate  $C_{\text{comp,sampled}}$  using the observation times through Q16 for the 40 hot Jupiter hosts with 2.8 days <  $P$  < 10 days in our sample (Table 1). In using the hot Jupiter hosts, we assume that their observational cadence is representative of that of proto-hot Jupiter hosts. The factor  $C_{\text{comp,sampled}}$  automatically incorporates  $C_{\text{comp}}$  (Equation (4)). We require that the planet transits three times. Accounting for missing data reduces  $r$  to 0.913, 0.668, 0.422 for  $P_{\text{final}} = 2.8, 5, 10$  days, respectively.

Second, we consider whether the transits have sufficient signal-to-noise to be detected. A signal-to-noise ratio (S/N) of 7 is the formal threshold for detection (Tenenbaum et al. 2012, 2013), but estimates of the current sensitivity of the *Kepler* pipeline vary; for example, Fressin et al. (2013) model the detection threshold as a ramp ranging from 0% at S/N of 8 to 100% at S/N of 16. For the expected progenitors of a given hot Jupiter, the signal-to-noise is (based on Howard et al. 2012, Equation (1)):

$$\text{S/N} = \frac{\delta}{\sigma_{\text{CDPP}}} \sqrt{N_{\text{transit}} \frac{t_{\text{dur}}}{t_{\text{CDPP}}}}, \quad (5)$$

for which  $N_{\text{transit}} = (t_{\text{survey}}(1 - e^2)^{3/2}/P_{\text{HJ}})$  is the average number of transits for a hot-Jupiter progenitor with  $P = P_{\text{HJ}}(1 - e^2)^{-3/2}$ ,  $t_{\text{dur}} = t_{\text{HJ}}/(1 + e \cos \omega)$  is the duration of the progenitor’s transit,  $\delta$  is the transit depth,  $\sigma_{\text{CDPP}}$  is the combined differential photometric precision (CDPP), and  $t_{\text{CDPP}}$



**Figure 2.** Top: relative probability of the predicted mean number of super-eccentric Jupiters (Interval 1: red dotted, Interval 2: blue dashed, total: black solid). Bottom: sampling from the above distribution of Poisson means to create a relative probability distribution of the expected number observed.

is the timescale of the CDPP. Therefore

$$\text{S/N} = \frac{\delta}{\sigma_{\text{CDPP}}} \sqrt{\frac{t_{\text{survey}} t_{\text{HJ}} (1 - e^2)^{3/2}}{P_{\text{HJ}} t_{\text{CDPP}} (1 + e \cos \omega)}}. \quad (6)$$

For each hot-Jupiter in our sample, we compute the most pessimistic S/N for a super-eccentric progenitor: a progenitor transiting at its periape with  $e = e_{\text{max}}$  three times over the duration of the survey. In each case, the S/N exceeds 22, which is well above the 100% detection threshold of 16 that was modeled by Fressin et al. (2013). Note that this approach automatically accounts for the effect of impact parameter on the transit duration by using the observed transit duration of the hot Jupiters. To get a sense for the *typical* S/N of a hot-Jupiter progenitor, we compute the signal-to-noise of a set of randomly generated progenitors, weighted by  $\dot{e}$ , the completeness (Equation (4)), and the transit probability. The resulting distribution of the S/N peaks at 100, with a 15th percentile value of 54 and 1 percentile value of 28. Therefore, we expect such progenitors to be readily detectable.

Next we derive the mean number of super-eccentric planets,  $\bar{N}_{\text{sup}}$ . To do so we insert the posteriors from Table 1 into Equation (1), making use of Equations (2) and (3), with the  $r$  listed above ( $r = 0.913, 0.668, 0.422$ ). We perform this procedure separately for Interval 1 and Interval 2, and use  $t_{\text{survey}} = 4$  yr, obtaining a  $\bar{N}_{\text{sup}}$  posterior for each interval that we sum to compute a total  $\bar{N}_{\text{sup}}$  (Figure 2). The total expected number is  $\bar{N}_{\text{sup}} = 5.3^{+2.4}_{-1.7}$ . This posterior represents a distribution of Poisson means.

We transform the distribution of means into a distribution of expected values by sampling  $N_{\text{sup}}$  from  $\bar{N}_{\text{sup}}$  according to Equation (B1). Each sample requires first drawing a mean ( $\bar{N}_{\text{sup}}$ ) from the distribution of means (Figure 2, top row) and then drawing an observed number  $N_{\text{sup}}$  from the Poisson distribution with that mean. The peak of the distribution is at 4, with a 1.3% chance of observing none.

We emphasize that after accounting for incompleteness, signal-to-noise, Poisson counting uncertainties, and updates to the observed samples, our final estimate is similar to the original S12 prediction of five to six super-eccentric Jupiters, assuming a completeness out to 2 yr orbital periods. However, given that



we find a lack of super-eccentric Jupiters, the considerations detailed in this section will prove essential to concluding that such planets are truly absent in nature, rather than just overlooked or coincidentally missing from this sample.

## 2.2. Prediction for Transit Light Curve Observables

We expect to be able to identify super-eccentric proto-hot Jupiters in the *Kepler* sample by fitting their transit light curves and identifying those for which the light curve model parameters are inconsistent with a circular orbit. A planet's orbital eccentricity affects its transit light curve in a number of ways (e.g., Barnes 2007; Ford et al. 2008; Kipping 2008). For long-period, highly eccentric, Jupiter-sized planets, the most detectable effect is on the transit duration. For a wide range of periape orientations relative to our line of sight, a planet on a highly eccentric orbit transits its star moving at a much larger speed than if it were on a circular orbit with the same orbital period. For Jupiter-sized planets, one can distinguish the effects of the transit speed on the ingress, egress, and full transit duration from the effects of the transit impact parameter and/or limb-darkening, even with long-cadence *Kepler* data (DJ12). Our fitting procedure fully incorporates all the uncertainty introduced by long exposure times.

For each planet, we fit a Mandel & Agol (2002) transit light curve model (binned to the exposure time, i.e., Kipping 2010a) with the following parameters: the planetary-to-stellar radius ratio  $R_p/R_*$ , the orbital period  $P$ , the inclination  $i$ , the scaled semi-major axis  $a/R_*$ , and the quadratic limb darkening parameters  $\mu_1$  and  $\mu_2$ . Assuming a circular orbit and a planetary mass much less than the stellar mass ( $M_p \ll M_*$ ), one can recast the scaled semi-major axis  $a/R_*$  (by substituting  $a = [GM_*(P/[2\pi])^2]^{1/3}$ ) as the bulk stellar density,  $\rho_* = M_*/(4/3\pi R_*^3)$ , which hereafter we will refer to as the stellar density. The stellar density measured from the light curve under the assumption of a circular orbit,  $\rho_{\text{circ}}$ , is related to the true stellar density,  $\rho_*$ , by:

$$\rho_*(e, \omega) g^3(e, \omega) = \rho_{\text{circ}}, \quad (7)$$

for which

$$g(e, \omega) = \frac{1 + e \sin \omega}{\sqrt{1 - e^2}} \quad (8)$$

is approximately the ratio of the observed transit speed (technically the transverse line-of-sight velocity) to the transit speed that the planet would have if it were on a circular orbit with the same orbital period (see Kipping 2010b and DJ12 for a detailed derivation). The argument of periape  $\omega$  represents the angle on the sky plane ( $\omega = 90^\circ$  for a planet transiting at periape).

We determine  $\rho_{\text{circ}}$  by fixing  $e = 0$ , allowing the stellar density to vary as a free parameter in the light curve model. The resulting  $\rho_{\text{circ}}$  is determined entirely by the shape and timing of the light curve. We then compare  $\rho_{\text{circ}}$  to the value of  $\rho_*$  determined through other methods (i.e., stellar models fit to the temperature and surface gravity determined through colors or spectroscopy). Although  $g$  is degenerate with the host star's density (Equation (7)), a loose (order-of-magnitude) constraint on  $\rho_*$  is sufficient for a tight constraint on the eccentricity (DJ12), the measurement of which we will describe and perform in Section 3. In the current section, we work with the variable  $\rho_{\text{circ}}/\rho_*$ . If  $\rho_{\text{circ}}/\rho_*$  is very large, then  $g$  must be large, and therefore the planet is moving more quickly during transits than a planet with the same orbital period  $P$  on a circular orbit. In

Appendix E, we summarize how our approach avoids problems caused by incorrect stellar parameters and approximations.

### 2.2.1. Expectations for Super-eccentric Planets

We perform a Monte Carlo simulation to predict the signature in the transit light curve observable  $\rho_{\text{circ}}/\rho_*$  expected from the super-eccentric proto-hot Jupiters (Section 2.1). We generate two-dimensional (2D) probability distributions in  $(P, \rho_{\text{circ}}/\rho_*)$  in Figure 3, where  $P$  is the orbital period, as follows:

1. We begin with an assumed  $P_{\text{final}}$ .
2. Using the completeness Equation (4), we generate a distribution of eccentricities  $\{e_i\}$  with a normalization constant  $C_{\text{norm}}$  following:

$$\text{Prob}(e) = \begin{cases} 0 & e > e_{\text{max}} \text{ or } e < 0.9 \\ C_{\text{norm}} C_{\text{comp, sampled}} |\dot{e}|^{-1} & 0.9 < e < e_{\text{max}} \end{cases} \quad (9)$$

3. For each eccentricity, we compute the corresponding orbital period  $P_i$  and randomly select an argument of periape  $\omega_i$ . Assuming a Sun-like star, we compute the scaled semi-major axis  $a_i/R_*$ .
4. We compute the transit probability:

$$\text{prob}_{\text{transit}} = \frac{R_*}{a_i} \frac{1 + e_i \sin \omega_i}{1 - e_i^2}. \quad (10)$$

Then we select a uniform random number between 0 and 1. If the number is less than the transit probability, we retain  $(e_i, \omega_i)$  in the distribution.

5. Then we compute  $\rho_{\text{circ}}/\rho_*$  using Equation (7).

We use the above procedure to generate four plots, corresponding to different  $P_{\text{final}}$  (Figure 3). In the fourth panel, instead of using a single  $P_{\text{final}}$ , we draw the  $P_{\text{final}}$  of each trial from the observed  $N_{P=P_{\text{final}}}$ , weighting each  $P_{\text{final}}$  by  $N_{\text{mod},0}/N_{P=P_{\text{final}},0}$  in the two intervals.

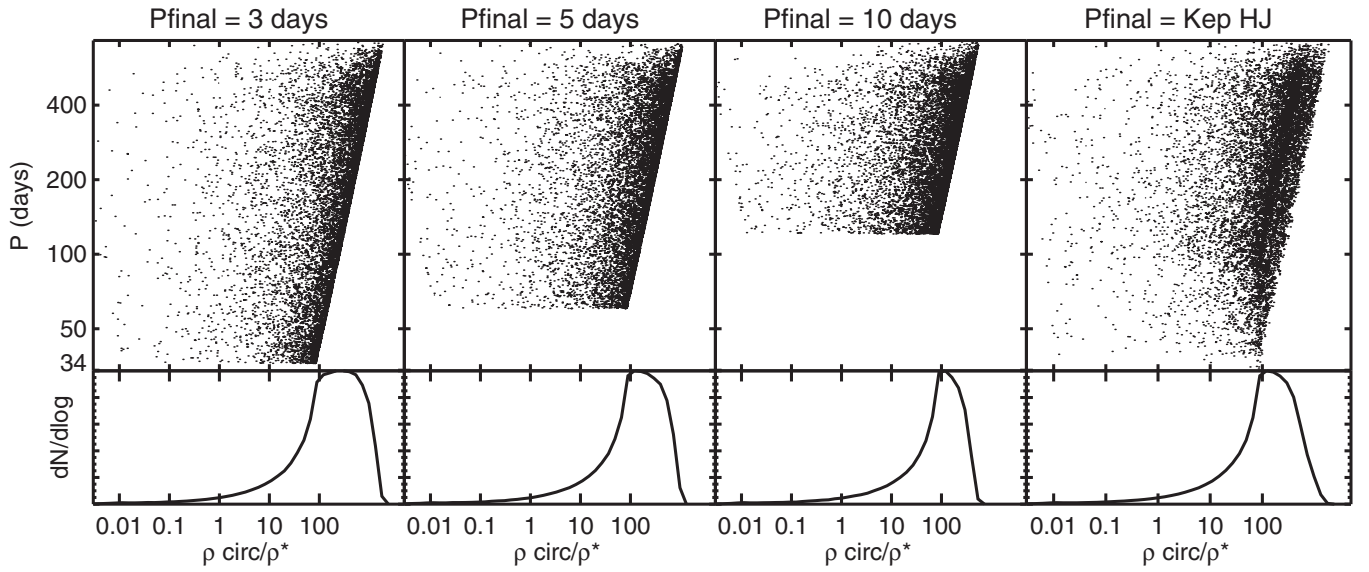
We see that a population of super-eccentric Jupiters will manifest itself as a collection of light curves with astrophysically implausible  $\rho_{\text{circ}}$  of 10–1000 times the estimated values for  $\rho_*$ . The super-eccentric proto-hot Jupiters will have orbital periods that range from  $P = 2.8 \text{ days}/(1-0.9^2)^{3/2} = 34 \text{ days}$  to 2 yr. About 90% of the expected planets have  $\rho_{\text{circ}}/\rho_* > 10$ , making them easy to identify.

### 2.2.2. Proto-hot Jupiters with $0.6 < e < 0.9$

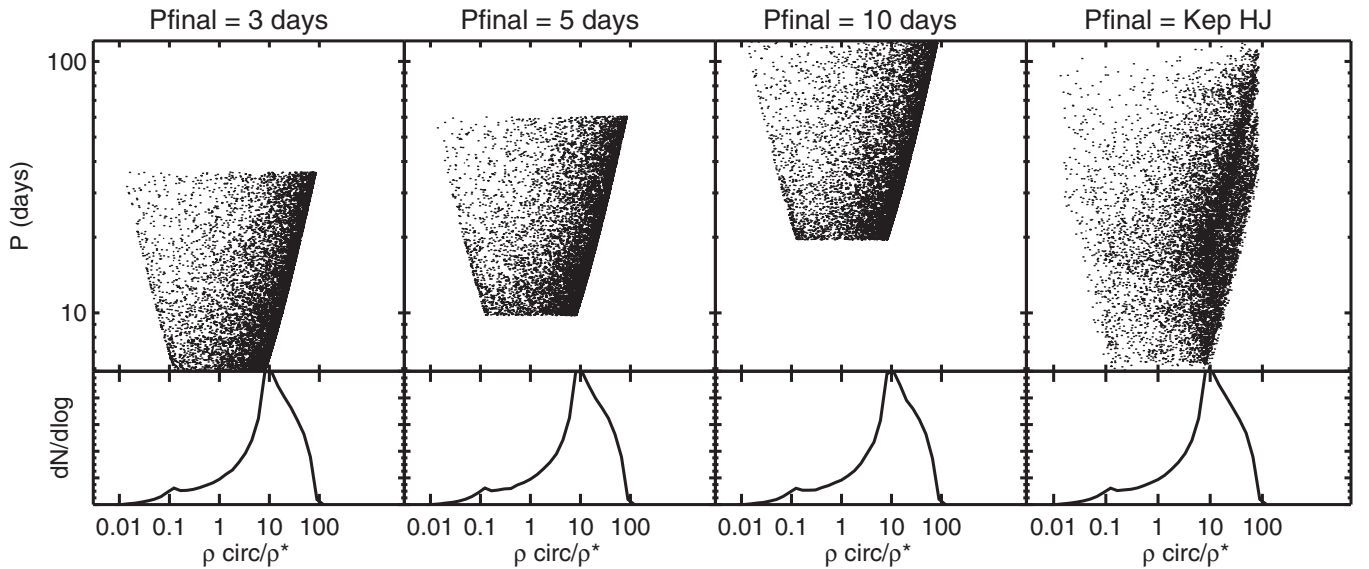
S12 focused their prediction on super-eccentric planets with  $e > 0.9$ . However, we also expect to find proto-hot Jupiters with less extreme eccentricities ( $0.6 < e < 0.9$ ) along the same  $P_{\text{final}}$  track. We repeat the procedure in Section 2.2.1 for the interval  $0.6 < e < 0.9$ . The overall occurrence rate for this interval is 0.61 relative to  $\bar{N}_{\text{mod}}$ . As shown in Figure 4, the proto-hot Jupiters in the  $0.6 < e < 0.9$  range have shorter orbital periods ( $6 < P < 121 \text{ days}$ ). However, their transit durations and the inferred stellar density from a circular fit are not as strikingly anomalous as for the super-eccentric proto-hot Jupiters, making them less easy to identify. Therefore, we do not focus on these objects, but discuss them further in the conclusion (Section 5).

## 3. RESULTS: A PAUCITY OF PROTO-HOT JUPITERS

We search for the super-eccentric proto-hot Jupiters predicted by S12 and find significantly fewer than expected. We describe our search procedure and present our measurements (Section 3.1) and assess the significance of this null result (Section 3.2).



**Figure 3.** Top: 2D posterior, orbital period  $P$  vs.  $\rho_{\text{circ}}/\rho_*$ , for planets with  $e > 0.9$  and  $P_{\text{final}} = 3, 5, 10$  days (panels 1–3) or  $P_{\text{final}}$  drawn from *Kepler* hot Jupiters with  $2.8 \text{ days} < P < 10 \text{ days}$  (panel 4). Bottom: Posterior  $\rho_{\text{circ}}/\rho_*$  marginalized over orbital period. Proto-hot Jupiters with  $e > 0.9$  should have anomalously large  $\rho_{\text{circ}}$  measured from the transit light curve compared to their estimated  $\rho_*$ , making them easy to identify. We expect approximately four super-eccentric proto-hot Jupiters in the high probability density region.



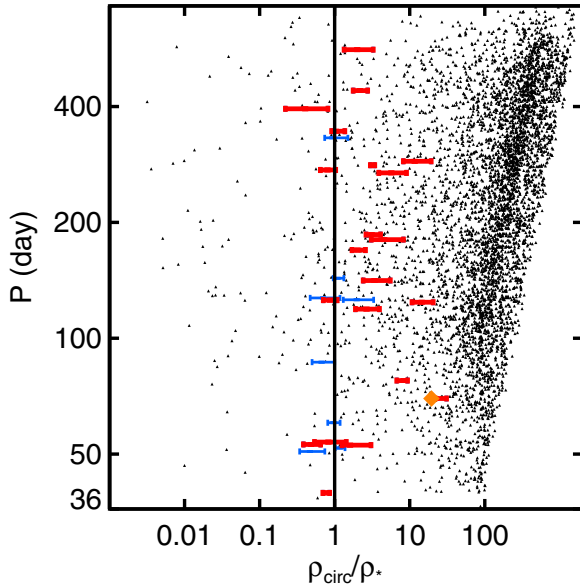
**Figure 4.** Top: 2D posterior for orbital period  $P$  vs.  $\rho_{\text{circ}}/\rho_*$  for planets with  $0.6 < e < 0.9$  and  $P_{\text{final}} = 3, 5, 10$  days (panels 1–3), or with  $P_{\text{final}}$  drawn from *Kepler* hot Jupiters in the interval  $2.8 \text{ days} < P < 10 \text{ days}$  (panel 4). Bottom: posterior  $\rho_{\text{circ}}/\rho_*$  marginalized over orbital period. Proto-hot Jupiters with  $0.6 < e < 0.9$  do not typically have such large  $\rho_{\text{circ}}/\rho_*$  as their super-eccentric ( $e > 0.9$ ) counterparts (Figure 3), making them less easy to identify.

### 3.1. Transit Light Curve Observables for Potential Proto-hot Jupiters

We begin by identifying planet candidates that conform to our selection criteria. Applying the same criteria as in Section 2, we identify candidates with  $8R_{\oplus} < R_p < 22R_{\oplus}$  and stellar parameters  $4500 \text{ K} < T_{\text{eff}} < 6500 \text{ K}$  and  $\log g > 4$  (or, for those with  $\log g < 4$ , consistent with  $\log g = 4$  within the uncertainty). We restrict the sample to candidates that have orbital periods between 34 days (corresponding to  $P_{\text{final}} = 2.8$  for  $e = 0.9$ ) and 2 yr, transit three times in Q1–Q16, have signal-to-noise above 10, and which do not blatantly fail false-positive vetting. See Appendix A for further details. We are left with 31 planet candidates, including Kepler-419-b (D12). None of selection criteria disfavor eccentric candidates, nor does the pipeline penalize planets for having a duration different than that

expected from a circular orbit. The consistency of our sample with Planet Hunters (discussed further in Appendix A), which were detected by eye, leads us to believe that the transits of long-period Jupiter-sized planets are not being missed due to transit timing variations. Moreover, Kepler-419-b—our most eccentric planet—was detected despite its 2 hr transit timing variations (TTVs).

For each candidate, we retrieve the publicly available data from MAST, use at least nine quarters of data (the amount available when we originally submitted this paper), and then supplemented some candidates with additional data as we revised the paper. We extract the transits using AutoKep (Gazak et al. 2012) and perform an MCMC fit using the Transit Analysis Package (TAP; Gazak et al. 2012). We fix  $e = 0$ , but allow all other parameters to vary, including noise parameters for the Carter & Winn (2009) wavelet likelihood function and



**Figure 5.** Expected 2D posterior for orbital period  $P$  vs.  $\rho_{\text{circ}}/\rho_*$  (taken from panel 4 of Figure 3). The values we measured for our 34 candidates are overplotted. Thin, blue bars: candidates with companions in their systems. Thick, red bars: candidates with no known companions. We do not see the expected approximately four candidates in the region of high-probability density. Orange diamond: Kepler-419-b (bottom).

first-order polynomial correction terms. We use short-cadence data when available and obtain each candidate’s  $\rho_{\text{circ}}$  posterior.

Next we follow<sup>10</sup> Section 3.3 of D12 to compute a  $\rho_*$  posterior for each host star using the Takeda et al. (2007) YREC stellar evolution models and the estimated effective temperature, surface gravity, and metallicity from Huber et al. (2014). We describe exceptions to this procedure in Appendix F.

Finally, we combine the  $\rho_{\text{circ}}$  and  $\rho_*$  posteriors into a posterior of  $\rho_{\text{circ}}/\rho_*$  for each candidate, marginalized over all other parameters. In Figure 5, we plot the resulting values on top of the probability distribution for the predicted super-eccentric proto-hot Jupiters (Figure 3, panel 4). None of the candidates fall in the high-probability area of the prediction. We indicate candidates with known companions in their system with blue bars; none can have  $e > 0.9$  and  $3 < P_{\text{final}} < 10$  days without its orbit crossing a companion’s. As expected, all candidates with companions have  $\rho_{\text{circ}}/\rho_*$  close to 1.

Three candidates (none of which have known companions) have  $\rho_{\text{circ}}/\rho_* > 10$ : KOI-211.01, Kepler-419-b (D12), and KOI-3801.01 (first discovered by Planet Hunters (Wang et al. 2013), and now a *Kepler* candidate). D12 found that Kepler-419-b has  $e = 0.81^{+0.10}_{-0.07}$  and  $P_{\text{final}} = 14^{+6}_{-10}$  days. The probability of it having  $e > 0.9$  and  $2.8 < P_{\text{final}} < 10$  days is 12%. Subsequently,  $e > 0.9$  was ruled out completely by our radial velocity follow up (Dawson et al. 2014), in which we precisely measured  $e = 0.83 \pm 0.01$ . Therefore we exclude this candidate from our statistical tests below. The probability of KOI-211.01 having  $e > 0.9$  and  $2.8 < P_{\text{final}} < 10$  days is 14%. In assessing the consistency of the observations with the prediction of Section 2, we will fully consider the possibility that KOI-211.01 might be a super-eccentric proto-hot Jupiter. However,

we note that it is on the list of eclipsing binaries (Slawson et al. 2011). Finally, the probability of KOI-3801.01 having  $e > 0.9$  and  $2.8 < P_{\text{final}} < 10$  days is 8%.

We emphasize that it is not surprising that many of the candidates lie in the low-probability region (likely Jupiters with  $e < 0.9$  or  $P_{\text{final}} > 10$  days, of which there may be any number). It is only surprising that we do not see approximately four in the high-probability region (given the assumptions in Appendix G).

### 3.2. Statistical Significance of Lack of Proto-hot Jupiters

None of the observed candidates lie in the high-probability-density region of Figure 5, so it is unlikely that the six predicted (Section 2.1) super-eccentric proto-hot Jupiters are present but missed. If we were certain that none of the candidates has  $e > 0.9$  and  $2.8 \text{ days} < P_{\text{final}} < 10 \text{ days}$ , the probability that observed number of super-eccentric proto-hot Jupiters agrees with the prediction would simply be 1.3%. (This is the probability, computed in Section 2, of observing zero super-eccentric proto-hot Jupiters given the Poisson uncertainties in the observed number of super-eccentric Jupiters and in numbers used to compute the prediction.) However, there is a small chance that there are indeed super-eccentric proto-hot-Jupiters among the sample, but that they just so happen to have their periaapses oriented in the narrow range of angles producing an unremarkable  $\rho_{\text{circ}}/\rho_*$ . Therefore we use a Monte Carlo procedure to assess the consistency of the  $\rho_{\text{circ}}/\rho_*$  posterior derived for each candidate with the predicted population of super-eccentric planets.

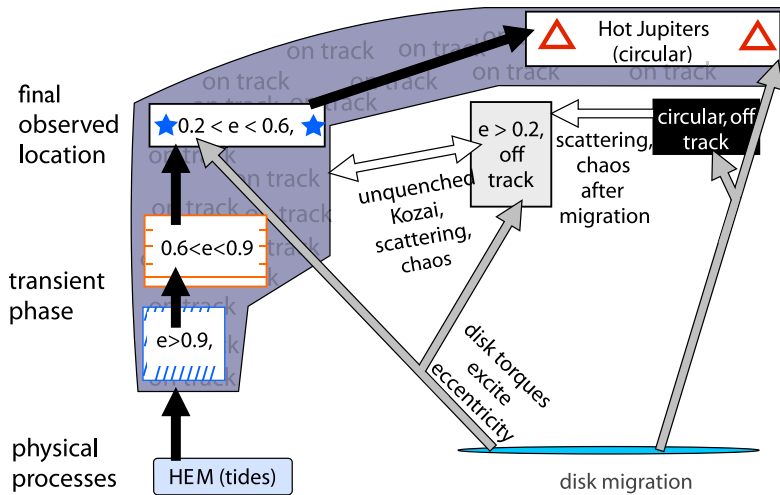
We first use the  $\rho_{\text{circ}}/\rho_*$  posteriors to generate an eccentricity posterior for each candidate via an MCMC exploration of a limited set of parameters:  $\rho_{\text{circ}}$ ,  $\rho_*$ ,  $e$ , and  $\omega$  (as outlined in DJ12, Section 3.4). Although we can only make a tight eccentricity measurement when the planet’s eccentricity is large (DJ12), the broad eccentricity posterior for the typical candidate here is useful for this purpose: it contains very little probability at the high eccentricities corresponding to  $e > 0.9$ ,  $2.8 \text{ days} < P_{\text{final}} < 10 \text{ days}$ . We describe exceptions to this procedure in Appendix F. Recently, Kipping (2014, K14 hereafter) presented several caveats for deriving  $e$  and  $\omega$  from  $\rho_{\text{circ}}$ . We discuss these caveats and why they are not an issue for this study in Appendix E.

We then perform  $10^6$  trials in which we randomly select an eccentricity from each candidate’s eccentricity posterior. We compute  $P_{\text{final}}$  and count  $N_{\text{sup}}$  in Intervals 1 and 2. If both are greater than or equal to the respective numbers drawn from posteriors in Figure 2, bottom panel (red dotted and blue dashed curves), we count the trial as a success, meaning that at least as many super-eccentric Jupiters as predicted were detected. 96.9% of trials were unsuccessful. We exclude the candidates with known companions from this procedure (Figure 5, thin blue bars), because it so happens that none of them can have  $e > 0.9$  and  $2.8 \text{ days} < P_{\text{final}} < 10 \text{ days}$  without crossing the orbit of another candidate in the system. We find that, with 96.9% confidence, we detected too few super-eccentric proto-hot Jupiters to be consistent with the prediction of Section 2. For example, 54% of trials had zero super-eccentric proto-hot Jupiters, 87% had one or fewer, and 98% had two or fewer. From these trials, we measure a  $N_{\text{sup}}$  posterior with a median  $N_{\text{sup}} = 0^{+1}_{-0}$  (89% confidence interval). No single planet is likely to be super-eccentric, but each has a small chance of being super-eccentric.

The Jeffrey’s prior we use in computing the posterior of the mean number of planets based on the observed number (Appendix B) has a conservative influence on our results. If

<sup>10</sup> Instead of imposing a prior on the stellar mass, metallicity, and age from a TRILEGAL (TRIdimensional model of the GALaxy; Girardi et al. 2005) synthetic *Kepler* field population, we assume a uniform prior on these model parameters, because a similar prior was already imposed by Huber et al. (2014) to generate the stellar parameters.





**Figure 6.** Schematic of pathways (arrows) for creating the observed giant planet populations, which we assume form at orbital periods beyond about 2 yr (corresponding to the uptick in giant planet frequency at around 1 AU, e.g., Cumming et al. 2008) and migrated closer to their stars via HEM or disk migration. On track (purple region) refers to the angular momentum range under consideration (i.e.,  $2.8 \text{ days} < P_{\text{final}} < 10 \text{ days}$ ). The black arrows indicate the path that we have assumed throughout the paper for HEM caused either by a planetary or stellar perturber. For example, a Jupiter may be perturbed by a stellar binary companion, follow the black arrow to the region of super-eccentric Jupiters with  $2.8 \text{ days} < P_{\text{final}} < 10 \text{ days}$  (blue, diagonal stripe region), undergo tidal circularization along its angular momentum track to  $e < 0.9$  (horizontal orange striped region), become a moderately eccentric Jupiter (blue stars), and eventually achieve hot-Jupiter-hood (red triangles). The other color arrows indicate alternative pathways caused by secular chaos, scattering, or unquenched Kozai moving the planets off track (white) or disk migration (gray), and the colors and patterns of the boxes correspond to the regions of parameter space in Figure 1. See the text for detailed discussion of the possible scenarios represented here.

we instead impose a uniform prior, the significance of the lack of super-eccentric Jupiters would be 97.6% instead of 96.9%. The significance of our results would decrease if we had a prior expectation against either super-eccentric Jupiters or the moderately eccentric Jupiters to which their occurrence is proportional, but we have no such expectations.

#### 4. EXPLAINING THE PAUCITY OF PROTO-HOT JUPITERS

So far (Sections 1–3) we have been considering a scenario in which hot Jupiters begin beyond orbital periods of 2 yr on super-eccentric orbits—caused by gravitational perturbations from a companion (e.g., star–planet Kozai, planet–planet Kozai, planet–planet scattering, secular chaos)—and subsequently undergo tidal circularization along a constant angular momentum track, reaching a final orbital period  $P_{\text{final}}$ . This process is known as HEM. We schematically summarize this (black arrows) and other possible origins for hot Jupiters (white and gray arrows), as well as moderately eccentric Jupiters with  $2.8 \text{ days} < P_{\text{final}} < 10 \text{ days}$ , in Figure 6. The corresponding populations from the non-*Kepler* sample (EOD; Wright et al. 2012) are plotted in Figure 1.

Now we relax previous assumptions about HEM and explore how we can account for the lack of super-eccentric proto-hot Jupiters (Section 3). In Sections 4.1 and 4.2, we relax the assumption that Jupiters began tidal circularization beyond an orbital period of 2 yr, finding that this possibility could indeed account for the lack of super-eccentric Jupiters. In Section 4.3, we consider the particular case of HEM via the Kozai mechanism in which the Kozai oscillations of the observed moderately eccentric Jupiters are unquenched. In Section 4.4, we relax the assumption of a steady current of hot Jupiters produced by HEM in the observed *Kepler* sample, but find that a lack of steady current is unlikely to account for the lack of super-eccentric proto-hot Jupiters. In Section 4.5, we place an upper limit on the fraction of hot Jupiters caused

by Kozai perturbations from a stellar binary companion. In Appendix G, we describe additional assumptions, including tidal assumptions, most of which we do not expect to affect our results.

##### 4.1. No Tidal Circularization: Hot Jupiters and Moderately Eccentric Jupiters Implanted Interior to 1 AU

Rather than starting on highly eccentric orbits with orbital periods above 2 yr, hot Jupiters and moderately eccentric Jupiters may have reached the region we observe today without tidal circularization. The moderately eccentric Jupiters (blue stars, Figures 1 and 6) observed along the angular momentum tracks have may have been placed there by whatever mechanism implanted eccentric Jupiters interior to 1 AU (gray region, 1). This population could possibly originate from disk migration. It remains debated whether planet-disk interactions could excite warm Jupiters’ eccentricities. Goldreich & Sari (2003) and Sari & Goldreich (2004) argued that disk migration can potentially excite moderate eccentricities through resonance torques. Recently Dunhill et al. (2013) modeled planet-disk interactions using high-resolution, three-dimensional simulations and found that disks are unlikely to excite the eccentricities of giant planets. However, Tsang (2014) and Tsang et al. (2014) made analytical arguments that planet eccentricities can be excited in a disks with entropy gradients. Planet–planet scattering directly from several AU is unlikely to be responsible for the moderately eccentric Jupiters, due to the large change in orbital energy required. Scattering following disk migration (e.g., Guillochon et al. 2011) is unlikely to produce sufficiently large eccentricities at such small semi-major axes because close encounters will lead to collisions rather than ejections (Ida et al. 2013; Petrovich et al. 2014). At such small semi-major axes, the relative velocities associated with a given eccentricity are large. Once giant planets stir each other to modest eccentricities, the relative velocities rival the escape velocities, and the cross section for collision becomes a significant fraction of the scattering cross section. Collisions damp the relative velocities, limiting



**Table 2**  
Occurrence<sup>a</sup> of Jupiters Detected by RV Surveys (Plotted in Figure 1)

Period range (days)	Eccentricity	Count	Poisson Range <sup>b</sup>	Number per log <sub>10</sub> <i>a</i>
2.8 < $P_{\text{final}}$ < 10	0.2 < $e$ < 0.6	4	4 ± 2	11 <sup>+6</sup> <sub>-5</sub>
10 < $P$ < 250	0.2 < $e$ < 0.6	18	18 <sup>+5</sup> <sub>-4</sub>	19 <sup>+5</sup> <sub>-4</sub>
<hr/>				
2.8 < $P$ < 5	0 < $e$ < 0.2	17	17 ± 4	100 <sup>+30</sup> <sub>-20</sub>
2.8 < $P$ < 10	0 < $e$ < 0.2	21	21 <sup>+5</sup> <sub>-4</sub>	57 <sup>+13</sup> <sub>-11</sub>
5 < $P$ < 10	0 < $e$ < 0.2	4	4.2 <sup>+2.4</sup> <sub>-1.7</sub>	21 <sup>+12</sup> <sub>-9</sub>
5 < $P$ < 250	0 < $e$ < 0.2	29	29 <sup>+6</sup> <sub>-5</sub>	26 <sup>+5</sup> <sub>-4</sub>
10 < $P$ < 250	0 < $e$ < 0.2	25	25 ± 5	27 <sup>+6</sup> <sub>-5</sub>

**Notes.**

<sup>a</sup> Numbers do not account for RV observational biases

<sup>b</sup> Range of Poisson means from whose distributions the count could have been drawn, computed using Jeffrey’s prior. See Appendix B for further details.

eccentricity growth. Petrovich et al. (2014) demonstrated that in-situ excitation cannot produce the observed moderate eccentricities of Jupiters within 0.15 AU (i.e., the blue stars in our Figure 1). However, secular chaos Wu & Lithwick (2011) following disk migration remains a possibility.

In Figure 1, the moderately eccentric Jupiters (blue stars) look as if they could be an extension of the distribution of warm Jupiters in the gray region. In Table 2, we compute the occurrence rate of giant planets ( $M_p \sin i > 0.25$ ) in the RV-discovered sample in different regions of Figure 1. The occurrence rate per log semi-major axis interval of moderately eccentric Jupiters with 2.8 days <  $P_{\text{final}}$  < 10 days is less than or equal to that in the gray region (10 days <  $P$  < 250 days). Therefore a separate mechanism for producing the blue stars apart from direct implantation may not be necessary. If non-tidal implantation was dominant, the number of moderately eccentric Jupiters should not be used to predict the number of super-eccentric Jupiters, because the moderately eccentric Jupiters did not tidally circularize from super-eccentric orbits.

If the moderately eccentric Jupiters with 2.8 days <  $P_{\text{final}}$  < 10 days did not undergo tidal circularization, hot Jupiters themselves could be part of a continuous distribution of circular Jupiters interior to 1 AU (Figures 1 and 6, black region), which must have migrated somehow. Disk migration effectively produces planets on circular orbits, but seems inconsistent with the high obliquities of hot Jupiters orbiting hot stars (Winn et al. 2010; Albrecht et al. 2012, but see also Rogers et al. 2012; Rogers & Lin 2013). However, disk migration may have produced some or all of the well-aligned hot Jupiters if their low obliquities are not the result of tidal realignment.

If the cut-off for tidal circularization is 2.8 days, rather than 10 days, Jupiters on circular orbits with  $P > 2.8$  days would actually be part of the so-called period valley, rather than the hot Jupiter pile-up. The period valley refers to the region exterior to hot Jupiters with  $P < 250$  days, where giant planets are scarce. The divide between hot Jupiters and the period valley (i.e., if it is 2.8 days, 10 days, or some other value) is ambiguous in the literature (e.g., Jones et al. 2003; Udry et al. 2003; Wright et al. 2009; Wittenmyer et al. 2010). The observed edge of hot Jupiters in ground-based transit surveys near 10 days may be partially caused by a combination of the reduced geometric transit probability of long-period planets and the inefficiency of ground-based transit surveys in detecting them (Gaudi et al. 2005), rather than a drop in the intrinsic occurrence rate. We note

that the distribution of giant planets inferred from the *Kepler* Mission, assessed out to 50 days (Youdin 2011; Howard et al. 2012), has no such edge. However, Wright et al. (2009) detect an edge at approximately 0.07 AU (five days) in their RV survey, which suffers from different (but less severe) biases than transit surveys. Thus, the existence and location of the cut-off remains uncertain.

The cut-off may in fact be *between*  $P = 2.8$  days and  $P = 10$  days. In Figure 1, the edge of the pile-up of circular Jupiters appears to end at around 0.057 AU (five days), as Wright et al. (2009) found. If we separate the hot Jupiters below this cut-off, we recover a pile-up of hot Jupiters: in the region from 2.8 days <  $P$  < 5 days we observe an excess of circular Jupiters that is inconsistent with the occurrence rate in the period valley by a factor of three (Table 2). If the cut-off for hot Jupiters is truly five days, the prediction for super-eccentric proto-hot Jupiters should be based only on the number of moderately eccentric Jupiters with 2.8 days <  $P_{\text{final}}$  < 5 days. In that case (repeating the calculations of Section 2.1), we expect to find only  $1 \pm 1$  super-eccentric proto-hot Jupiters with  $e > 0.9$  and 2.8 days <  $P_{\text{final}}$  < 5 days (72% confidence interval), and our confidence that we found fewer than predicted (Section 3.2) drops to 72%.

#### 4.2. Some or All Proto-hot Jupiters May Have Bypassed the $e > 0.9$ Portion of the $P_{\text{final}}$ Track

Alternatively, the typical hot Jupiter may have undergone tidal circularization but bypassed the high eccentricity phase, starting on the HEM track with  $0.6 < e < 0.9$  in the region indicated by orange stripes in Figures 1 and 6 (or even in the  $0.2 < e < 0.6$  region). For  $P_{\text{final}} < 10$  days, a Jupiter would begin on the HEM track at an orbital period less than 120 days, or 0.5 AU around a Sun-like star. The Jupiter is unlikely to have formed here—the critical core mass required to accrete a massive atmosphere most likely exceeds the amount of refractory materials available (Rafikov 2006; but see also Lee et al. 2014)—or have been scattered here from several AU, which would require a large change in orbital energy.

However, the typical hot Jupiter may have been delivered to this region by disk migration followed by eccentricity excitation (Guillochon et al. 2011) through secular chaos (Wu & Lithwick 2011) or planet–planet scattering. However, in situ excitation by planet–planet scattering to the eccentricity values in the orange striped region may not be feasible because the corresponding random velocity,  $v_{\text{rand}} \sim e\sqrt{GM_*/a}$  is a significant fraction of a giant planet’s escape velocity,  $v_{\text{esc}} = \sqrt{2GM_p/R}$ . In this super-Hill-velocity regime, as long as  $v_{\text{rand}} < v_{\text{esc}}$ , the cross section for collisions is of order  $(v_{\text{rand}}/v_{\text{esc}})^2$  times the cross section for stirring (e.g., Goldreich et al. 2004). For  $v_{\text{rand}} > v_{\text{esc}}$ , gravitational focusing is ineffective and only collisions occur. Since  $v_{\text{rand}}/v_{\text{esc}} \sim e(a/0.2 \text{ AU})^{-1/2}(M_p/M_J)^{-1/2}(R_p/R_J)^{1/2}$ ,  $v_{\text{rand}} \rightarrow v_{\text{esc}}$  in the vicinity of the orange striped region, and planet–planet interactions may be dominated by collisions rather than eccentricity growth. See Petrovich et al. 2014 for an exploration of this phenomenon at smaller semi-major axes and Ida et al. (2013), Figure 7 and Section 5.3, and Jurić & Tremaine (2008), Figure 8, for scattering simulations in which the resulting eccentricities are lower for giant planets with  $a < 1$  AU.

The overall picture of this scenario is that proto-hot Jupiters start on the HEM track inside 1 AU with eccentricities similar to those of the planets we observe in the period valley, rather than starting with  $e \rightarrow 1$  at large semi-major axes. Assuming

a steady flux of proto-hot Jupiters into the orange striped region, if the four moderately eccentric, RV-detected<sup>11</sup> Jupiters (corresponding to  $\bar{N}_{\text{mod}} = 4 \pm 2$ ) in Figure 1 (blue stars) originated from the orange region, we would expect

$$\frac{\bar{N}_{\text{mod}} \int_{0.6}^{0.9} C_{\text{comp}}(e) |\dot{e}|^{-1} de}{\int_{0.2}^{0.6} C_{\text{comp}}(e) |\dot{e}|^{-1} de} = 4 \pm 2 \times 0.612 = 2.6^{+1.5}_{-1.0}$$

proto-hot Jupiters in the orange region. We see one such planet, HD 17156 b (Figure 1); the probability of observing only one such planet is 5%. Therefore we also observe a paucity of eccentric Jupiters in the orange region, which means this is not a satisfactory explanation.

Since we observe HD 80606 b in the blue striped region, all proto-hot Jupiters must not begin in the orange striped region. Secular chaos or planet–planet scattering, taking place either where the planet formed or disk-migrated to, may place proto-hot Jupiters in both the orange striped region and the blue striped region, with the majority in orange striped region. The proto-hot Jupiters in these two regions would be created by the same dynamical processes responsible for Jupiters with  $P_{\text{final}} > 10$  days (failed hot Jupiters), of which we observe more with  $0.6 < e < 0.9$  than with  $e > 0.9$  (although this may be partly due to observational bias).

#### 4.3. Proto-hot Jupiters Created by Planet–Planet Kozai

If a proto-hot Jupiter’s Kozai oscillations have not been quenched at the time of observation, we may observe the proto-hot Jupiter off its tidal-circularization track. The possibility that super-eccentric proto-hot Jupiters spend time at low eccentricities does not reduce their expected number. S12 demonstrate that tidal dissipation primarily occurs during high-eccentricity intervals. Regardless of how much time the Jupiter spends off its  $a_{\text{final}}$  track during low-eccentricity Kozai phases, it spends the same total amount of time on the  $a_{\text{final}}$  track undergoing tidal dissipation. Therefore Equation (1) predicts the total number of super-eccentric Jupiters *observed on the track*.

However, it is possible that the Kozai oscillations of the moderately eccentric calibration proto-hot Jupiters ( $N_{\text{mod},0}$  blue stars in Figure 1) are not quenched. We may be observing them in the low-eccentricity portions of their cycles, and they may actually be tidally dissipating on a track with  $P_{\text{final}} < 2.8$  days. If so, they should not have been used to compute the expected number of super-eccentric proto-hot Jupiters with  $2.8 < P_{\text{final}} < 10$  days. We clarify that even if the moderately eccentric Jupiters periodically oscillate to higher eccentricities corresponding to a dissipation track with  $P_{\text{final}} < 2.8$  days, they should not be used to compute the number of super-eccentric Jupiters with  $P_{\text{final}} < 2.8$  days, because Equation (1) only applies to planets observed on the track. Supporting this possibility, while hot Jupiters do not have nearby companions (Steffen et al. 2012), Dong et al. (2014) find that of the sample of warm Jupiters with known companions, all the companions are sufficiently massive and nearby for the Kozai cycles to remain unquenched. Additionally, several of the moderately eccentric Jupiters (HAT-P-34 b, HAT-P-31 b, and WASP-8 b) have linear trends in

the RV observations (Wright et al. 2011), indicating the presence of a companion.

We rewrite the condition for Kozai cycles to not be quenched from S12 (Equation (12)) as (assuming a Sun-like star and circular perturber with a minimum mutual inclination near  $40^\circ$ ):

$$a_{\text{per}} \leq 15 \text{ AU} \left( \frac{1 - 0.32^2}{1 - e^2} \right) \left( \frac{a_{\text{final}}}{0.05 \text{ AU}} \right)^{8/7} \left( \frac{M_{\text{per}}}{M_\odot} \right)^{2/7}, \quad (11)$$

for which  $a_{\text{per}}$  and  $M_{\text{per}}$  are the semi-major axis and mass of the putative perturbing companion, and  $a_{\text{final}}$  and  $e$  are the quantities for the moderately eccentric proto-hot Jupiter at the peak eccentricity of its Kozai cycle. All the warm Jupiters in our sample have been observed via RV observations, and such a nearby star would be evident via a dual set of lines or major inconsistency in the RV forward-modeling process. Therefore the perturber must be a planet, so hereafter we will refer to the possibility discussed in this section as planet–planet Kozai.

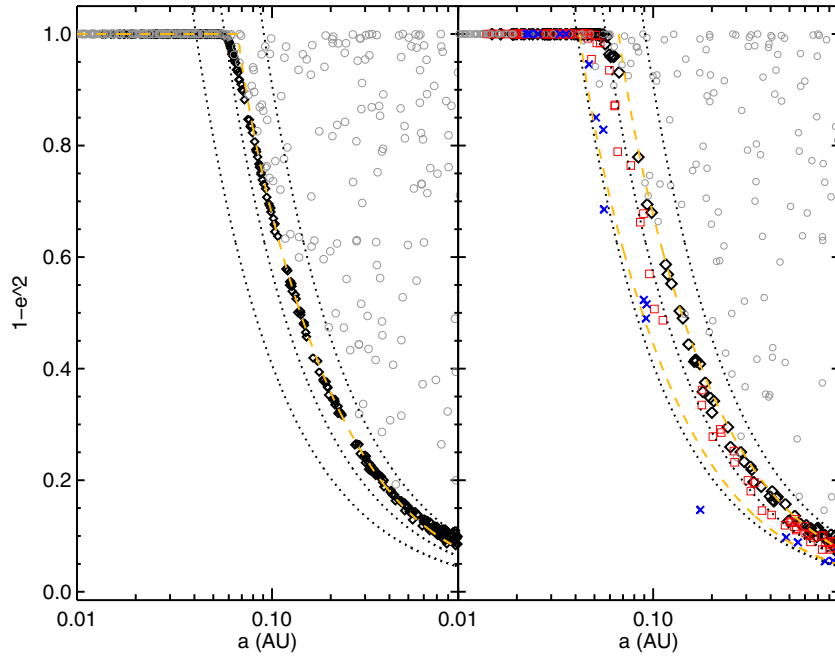
#### 4.4. Alternatives to the Steady Current Approximation

The S12 prediction of a readily observable number of super-eccentric proto-hot Jupiters assumed a steady current of proto-hot Jupiter production based on the fact that stars have been steadily produced throughout the Galaxy. In a sample of stars of identical ages, we would expect a steady current of proto-hot Jupiters only if the rate of hot Jupiter production throughout a star’s lifetime is constant (e.g., that a hot Jupiter is just as likely to be produced between 4.1 and 4.2 Gyr as it is during the first 100 Myr). This seems unlikely. In the HEM mechanisms proposed (Section 1), proto-hot Jupiters are spawned (i.e., begin their HEM journey) on instability timescales (planet–planet scattering, secular chaos) or the Kozai timescale, which are unlikely to always coincide with the typical stellar lifetime. More likely, the distribution of timescales is uniform (or normal) in order of magnitude and thus most proto-hot Jupiters are spawned early in their host stars’ lifetimes. Indeed, Quinn et al. (2012) recently discovered hot Jupiters in the 600 Myr Beehive cluster, and found that, accounting for the cluster’s enhanced metallicity, the hot Jupiter occurrence rate is consistent with that of the solar neighborhood.

In a sample of stars with uniform random ages spanning zero to the age of the Galaxy, we would expect a steady current, even if most proto-hot Jupiters are spawned early in their host stars’ lifetime, as long as the conditions for forming hot Jupiters have not changed over time. In practice, we expect young stars to be rotating too rapidly to be amenable to Doppler observations and too uncommon in our stellar neighborhood to make up a representative sample in transit surveys. Therefore, because of selection bias, the steady current approximation is unlikely to exactly hold.

Even if all the proto-hot Jupiters in our sample were spawned simultaneously, *we would still expect to observe proto-hot Jupiters*. However, they would be restricted to the narrow range of  $a_{\text{final}}$  tracks for which the circularization timescale is of order a star’s age, instead of being found along all  $a_{\text{final}}$  tracks in proportion to the circularization timescale ( $e/\dot{e} \propto a_{\text{final}}^8$ ). Inspired by population simulations by Hansen (2010, 2012), we simulate an extreme scenario in which every proto-hot Jupiter in the observable sample is created simultaneously (Figure 7, left panel). We begin with a population of Jupiters that is uniformly distributed in eccentricity and semi-major axis, extending to 10 AU; gray open circles had initial semi-major axes interior to

<sup>11</sup> In this calculation, we use the RV-detected sample. Even though transit probability is constant along an  $a_{\text{final}}$  track, ground-based transit surveys are still strongly biased against detecting planets transiting with longer orbital periods. RV samples suffer from their own biases against long period and eccentric planets, which we do not account for here.



**Figure 7.** Monte Carlo simulation of tidal evolution of proto-hot Jupiters assuming that all the proto-hot Jupiters were created at once (left panel) or that there is a steady current (right panel). In each simulation, proto-hot Jupiters are drawn from a distribution uniform in eccentricity and semi-major axis, extending to 10 AU. Planets that began interior to 1 AU are marked as open, gray circles. We overplot tracks of constant angular momentum (dotted lines) corresponding to  $P_{\text{final}} = 3, 5, 10$  days, as well as lines defined by a constant orbital change timescale (dashed orange lines, Equation (12)). In the right panel, the outer orange-dashed line represents a timescale that is 40 times longer than the inner orange dashed line. The red squares (blue x) were created 2 (13) times more recently than the black diamonds.

1 AU (representing the possibility that planets can begin HEM interior to 1 AU, as discussed in Section 4.2). In reality, the initial conditions will be set by whatever dynamical processes excite the planet’s eccentricity from the circular orbit it formed on. Since here we are agnostic about which of the proposed dynamical mechanisms is at play for initially exciting the proto-hot Jupiter’s eccentricity, we use a uniform distribution to get a sense of the evolution.

Then we evolve the tidal evolution equations until Jupiters with  $P_{\text{final}} < 5$  days have circularized. This is not absolute timescale, but relative to the unknown tidal dissipation constant. In Figure 7 we overplot tracks of constant angular momentum (dotted lines), as well as lines defined by a constant orbital change timescale set equal to the time of the plotted snapshot,

$$t_{\text{move}} = [(\dot{a}/a)^2 + (\dot{e}/e)^2]^{-1/2}, \quad (12)$$

(dashed orange lines) which match the  $a_{\text{final}}$  tracks at low eccentricities. This relationship represents an envelope in  $(a, e)$  space for a population of planets undergoing migration over a particular timescale ( $t_{\text{move}}$ ), rather than the evolution of a particular planet along a track. Although there is no steady current, we see a track consisting of (1) Jupiters along the same  $a_{\text{final}}$  track but with different starting eccentricities/semi-major axes, and (2) Jupiters along close, adjacent  $a_{\text{final}}$  tracks (those along the slightly larger  $a_{\text{final}}$  track have higher eccentricities because  $e/\dot{e} \propto a_{\text{final}}^8$ ).

For comparison (Figure 7, right panel) we perform a simulation featuring a steady current of hot Jupiters. In this case, the proto-hot Jupiters are distributed over a range of angular momenta tracks, but are most common (black diamonds) along the track where the tidal circularization time is order the total elapsed time (i.e., most of the proto-hot Jupiters began their HEM early in the lifetime of the oldest stars in the

sample). The smaller  $a_{\text{final}}$  tracks (red squares, blue X) are more sparsely populated because these planets circularize very quickly and we happen to be catching some. The left and right panels are not strikingly different. Particularly with a small observational sample size, we are unlikely to be able to distinguish whether we are seeing a narrow range of  $a_{\text{final}}$  tracks due to a lack of steady current (left) or simply due to a higher relative population along the  $a_{\text{final}}$  track of order a stellar age (right).

However, without a steady current of hot Jupiters, the number of moderately eccentric Jupiters in a  $P_{\text{final}}$  range cannot be used in Equation (3) to predict the number of super-eccentric Jupiters. The number of super-eccentric Jupiters would depend on the initial conditions generated by HEM mechanisms: the relative number of hot Jupiters along adjacent angular momentum tracks and beginning at different semi-major axes along the same track. However, if the initial eccentricities are roughly independent of semi-major axis, the distribution today would resemble one resulting from a steady current. The observed lack of super-eccentric proto-hot Jupiters would require very fine-tuned initial conditions, such as substantially fewer Jupiters beginning along a slightly larger angular momentum track. Therefore the paucity we found in Section 3 is unlikely to be fully accounted for by a lack of steady current.

Throughout this work we assumed that proto-hot Jupiters *can* travel along HEM tracks with circularization timescales of order a stellar lifetime. However, a nearby planetary perturber can potentially permanently remove the proto-hot Jupiter from the angular momentum track before tidal circularization decouples it. For example, in secular chaos a planet may be perturbed to a high eccentricity, begin to circularize along a track with a timescale longer than the chaos timescale, but then be chaotically perturbed by nearby planets again. If all proto-hot Jupiters are created early in the star’s lifetime and only those



with extremely short tidal circularization timescales escape the perturbations of nearby planets, we would indeed see a lack of super-eccentric proto-hot Jupiters in a sample that lacks young stars. In this framework, the observed moderately eccentric Jupiters would not have been produced by HEM, but by some other mechanism, as explored in Section 4.1. Their survival indicates that the circularization timescale is not extremely short in  $2.8 \text{ days} < P_{\text{final}} < 10 \text{ days}$ , and therefore the argument that planets can only travel along paths with very short circularization timescales would only apply for hot Jupiters with  $P < 2.8 \text{ days}$ .

Finally, a related issue is whether the ratio of moderately eccentric Jupiters to Jupiters with  $P = P_{\text{final}}$  is the same in the *Kepler* and calibration samples (i.e.,  $\bar{N}_{\text{mod}}/\bar{N}_{P=P_{\text{final}}} \simeq \bar{N}_{\text{mod},0}/\bar{N}_{P=P_{\text{final},0}}$ ), as assumed in Equation (2). We would expect the pile-up of hot Jupiters to be greater in the older sample (i.e., the *Kepler* sample) and therefore that Equation (2) might over-predict the number of super-eccentric Jupiters. However, we note that the overall occurrence rate of hot Jupiters in the *Kepler* sample is actually *smaller* than the in the RV-sample (Howard et al. 2012; Wright et al. 2012), so this is unlikely to be a problem in practice.

#### 4.5. Upper Limit on Star–Planet Kozai Contribution

If the observed moderately eccentric proto-hot Jupiters began with orbital periods above 2 yr and with  $e \rightarrow 1$ , and underwent tidal circularization while staying on a track of constant angular momentum (black arrows, Figure 6), we would expect to see four super-eccentric proto-hot Jupiters (Section 2); the lack of such planets indicates that one or more of the alternative pathways in Figure 6 (white and gray arrows) may dominate. These alternative pathways originate from a planetary perturber or disk, rather than a stellar perturber (see Section 4.3 for a justification of why a planetary Kozai perturber would be required). Here we place an upper limit on the fraction of hot Jupiters that followed the black arrow channel of HEM, beginning with a super eccentricity and moving along a track of constant angular momentum. Since this is the only pathway open to hot Jupiters produced by star–planet Kozai oscillations, the upper limit is also on the fraction of hot Jupiters created by stellar binaries. We repeat the MCMC procedure in Section 3.2 but update Equation (1) with an additional parameter,  $f_*$ , representing the fraction of hot Jupiters that undergo HEM from super-eccentricities (or, equivalently, the maximum fraction produced by star–planet Kozai):

$$\bar{N}_{\text{sup}} = f_* r(e_{\text{max}}) \bar{N}_{\text{mod}}. \quad (13)$$

We impose a modified Jeffrey’s prior on  $f_*$ , so the prior is uniform between 0 and 5% and scales with  $1/f_*$  above  $f_* = 5\%$ . We obtain a two-sigma limit of 44% on the fractional contribution from stellar binaries. Therefore, we expect that at least half of hot Jupiters were created by a planetary perturber (or a disk). We note that this result technically only applies to hot Jupiters with  $P > 2.8 \text{ days}$ , because no super-eccentric Jupiters with  $P_{\text{final}} < 2.8 \text{ days}$  were expected. This limit is consistent with a 30% contribution rate found by Naoz et al. (2012) and the  $<20\%$  contribution rate found by Petrovich (2014). This limit also implies that we would need at least a 44% false positive rate to account for the discrepancy. As discussed in Appendix G, such a high false positive rate is unlikely based on previous studies, and moreover, our additional vetting (Appendix A) results in our sample containing fewer false positives than the *Kepler* candidate list.

## 5. CONCLUSION

S12 predicted that if high-eccentricity migration (HEM) is the primary channel for producing hot Jupiters, the *Kepler* candidate collection should harbor a population of super-eccentric Jupiter-sized planets that are in the midst of tidal circularization. We developed and implemented a procedure to use the publicly available *Kepler* transit light curves to test this prediction and found a paucity of proto-hot Jupiters on super-eccentric orbits. Incorporating uncertainties due to counting statistics, uncertainties in the fitted light curve parameters and in the assumed stellar parameters, incompleteness due to the limited observational timespan and missing data, and the signal-to-noise limit, we expected to observe  $N_{\text{sup}} = 4$  (Section 2) but instead found only  $0^{+1}_{-0}$  (Section 3). (Note that we did not find one Jupiter that is likely to be super-eccentric; rather each Jupiter has a small chance of being super-eccentric, but with its orbit oriented so that its transit speed is unremarkable.) False positives are not likely to account for the discrepancy. The lack of super-eccentric proto-hot Jupiters may indicate that the assumed constant tidal time lag approximation—which sets the ratio of super-eccentric proto-hot Jupiters to the observed, partially circularized, moderately eccentric Jupiters used to compute the prediction—is incorrect (Appendix G). However, violation of this assumption could only account for the discrepancy if tidal dissipation were actually much *stronger* at high eccentricities along a given angular momentum track. The statistical significance of our results is 96.9%. In the future, we will use a pipeline targeted to finding large-period, Jupiter-sized planets (including those with only two transits, allowing us to extend the maximum orbital period considered from two years to four years) and inject the transits of simulated super-eccentric Jupiters to ensure that the assumptions about completeness we have made here are correct.

In Section 4, we explored a number of dynamical explanations for the paucity of super-eccentric proto-hot Jupiters, relaxing the assumptions by S12 that proto-hot Jupiters begin HEM beyond a 2 yr orbital period and that a steady current of hot Jupiters is being produced in the observed *Kepler* sample. We found that the lack of super-eccentric planets could be explained by one of the following scenarios. First, hot Jupiters with  $P > 2.8 \text{ days}$  may be implanted interior to 1 AU, and only those with  $P < 2.8 \text{ days}$  have undergone tidal circularization. This would be the case if either the tidal circularization timescale is typically only less than a stellar lifetime for  $P_{\text{final}} < 2.8 \text{ days}$ , or only proto-hot Jupiters with fast circularization timescales can manage to complete their circularization without being moved by a nearby planetary perturber. Second, the moderately eccentric Jupiters used to calibrate the prediction may be undergoing Kozai eccentricity oscillations caused by a nearby planetary perturber and we are observing them in the low-eccentricity phase, in which they are not currently experiencing tidal dissipation. In that case, we would not expect to observe any super-eccentric Jupiters currently experiencing tidal dissipation.

All these explanations point either to disk migration or to secular chaos, planet–planet scattering, or planet–planet Kozai (or other yet-to-be-proposed dynamical mechanism) as the dominant channel for hot-Jupiter production, rather than the star–planet Kozai mechanism. In Section 4, we placed an upper limit of 44% on the contribution of star–planet Kozai to hot Jupiters, which is consistent with the findings of Naoz et al. (2012). Our limit only applies to hot Jupiters with orbital periods greater than 2.8 days, as the prediction for super-eccentric Jupiters only applied to those ending their HEM journey at  $2.8 \text{ days} < P < 10 \text{ days}$ .

In this paper, we explored S12’s prediction for proto-hot Jupiters, but they made a similar prediction for a population super-eccentric binary stars, which they subsequently discovered (Dong et al. 2013). It would not be surprising if short-period stars were produced by the Kozai mechanism, but short-period planets primarily by scattering and chaos, which can potentially deliver the planets observed interior to 1 AU without the planets undergoing a super-eccentric phase. The initial conditions for stellar systems and planetary systems may differ in that planetary systems are both theorized (e.g., Barnes & Raymond 2004) and observed (e.g., Wright et al. 2009; Lissauer et al. 2011; Mayor et al. 2011) to form packed with many planets, which is a condition that may often lead to scattering and secular chaos. In contrast, stellar multiples are typically hierarchical, which is an optimal setup for the Kozai mechanism.

The lack of super-eccentric proto-hot Jupiters is a new piece of evidence that the models for the origins of hot Jupiters’ close-in orbits must match, joining the distribution of spin orbit measurements (e.g., Fabrycky & Winn 2009; Morton & Johnson 2011a; Naoz et al. 2012). We recommend that future theoretical studies of dynamical models for forming hot Jupiters predict the distribution of timescales for instabilities that form proto-hot Jupiters, how often the high-eccentricity phase of HEM is bypassed, the initial conditions along the HEM angular momentum tracks, and the expectations for high-eccentricity failed hot Jupiters that likely have periapses too distant to undergo tidal circularization, such as Kepler-419-b (D12). For the brightest *Kepler* host stars, we recommend measuring the spin-orbit alignment of planets in the period valley, whose obliquities have presumably not been affected by tides. Such measurements could elucidate whether the planets in the period valley have a single origin, or if there are two populations that might correspond to the circular planets and the eccentric planets. Additionally, we recommend investigating whether a gas disk could flatten and circularize a period valley planet’s orbit if the planet were scattered there before the gas disk dissipated.

We recommend that observers strive to better characterize the eccentricity distribution of the period valley, which we argued may be the launching point for the typical hot Jupiter’s HEM journey. It would be helpful to assess if the occurrence rate of eccentric Jupiters in this region is—when extrapolated to the 3 days  $< P_{\text{final}} < 10$  days region—sufficient to launch all the hot Jupiters interior to 1 AU. We also recommend that observers attempt to pin down the period or semi-major axis cut-off between the hot Jupiter pile-up and the period valley. Finally, although we found that it would be more difficult to identify proto-hot Jupiters with  $0.6 < e < 0.9$  using the photoeccentric effect, it could be feasible with more accurate and precise stellar parameters. We recommend spectroscopic follow-up of KOI host stars for this purpose.

We are thankful to the anonymous referee(s) for an insightful report that improved the paper, and in particular for advocating more conservative assumptions about the completeness of the *Kepler* pipeline. R.I.D. gratefully acknowledges the National Science Foundation Graduate Research Fellowship under grant DGE-1144152 and the Miller Institute for Basic Research in Science, University of California, Berkeley. J.A.J. is grateful for the generous grant support provided by the Alfred P. Sloan and David & Lucile Packard foundations. We are grateful to Smadar Naoz for many enlightening discussions and comments, including opening our eyes to other possibilities in

Section 4, for which we also thank Simon Albrecht and Fred Rasio. Many thanks to Adrian Barker, Rick Greenberg, Renu Malhotra, Francesca Valsecchi, and especially Brad Hansen for tidal insights; to Katherine Deck, Will Farr, Vicky Kalogera, Yoram Lithwick, and Matthew Payne for helpful dynamical discussions; to Will Farr and Moritz Günther for helpful statistical discussions; to Courtney Dressing and Francois Fressin for *Kepler* assistance; and to Joel Hartmann for helpful discussions about the ground-based sample. We are grateful to Subo Dong for constructive comments on a manuscript draft. Thanks to Joshua Carter, Boas Katz, Doug Lin, Geoff Marcy, Darin Ragozzine, Kevin Schlaufman, and Aristotle Socrates for helpful comments, and to Thomas Barclay, Christophe Burke, Jon Jenkins, and Jason Rowe for helpful discussions of *Kepler*’s completeness for giant planets. We benefitted from helpful conversations with Thomas Barclay about candidate vetting and Howard Isaacson and Rea Kolb about false positives. We are very grateful to Chelsea Huang for helpful discussions and for providing us with a detrended light curve for KIC 6805414. R.I.D. thanks David Charbonneau, Sean Andrews, Debra Fischer, Matt Holman, and Abraham Loeb for helpful comments on the thesis chapter version of this manuscript. Special thanks to J. Zachary Gazak for helpful modifications to the TAP code. We thank Dáithí Stone for making his library of IDL routines available.

This paper includes data collected by the *Kepler* mission. Funding for the *Kepler* mission is provided by the NASA Science Mission directorate. We are grateful to *Kepler* team for their work in making this revolutionizing mission possible and making the rich *Kepler* data set available. Some of the data presented in this paper were obtained from the Multimission Archive at the Space Telescope Science Institute (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NNX09AF08G and by other grants and contracts. This research has made use of the Exoplanet Orbit Database, and the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program.

## APPENDIX A

### SAMPLE OF LONG-PERIOD *KEPLER* GIANT PLANET CANDIDATES

We check for consistency among several sources to ensure that our sample is complete for planets that transit three times in Q1–Q12.

1. The Batalha et al. (2013a) candidate sample, which is complete for Jupiter-sized planets that transit three times within Q1–Q6, and also contains a number of planets that transit one or more times in Q1–Q8, with unknown completeness.
2. The Q1–Q12 threshold-crossing events table (TCE; acquired from the NASA Exoplanet Archive; see Tenenbaum et al. 2012; Tenenbaum et al. 2013), which is complete for Jupiter-sized planets that transit three times within Q1–Q12 ( $t_{\text{survey}} = 3$  yr), but the objects have not necessarily been vetted as candidates. We compile a list of TCE that meet our stellar and planetary criteria and remove those that are obviously instrument or stellar variability artifacts, or exhibit a deep secondary eclipse (indicating that they are eclipsing binaries). We use the data validation reports available for

- each TCE at the Exoplanet Archive to also exclude TCE with different odd and even transit depths (indicating that they are eclipsing binaries) or centroids that strongly correlate with the transits (indicating a blend). In total we find four candidates not present in the Batalha et al. (2013a) sample: KIC 12735740 (orbital period 282 days, now KOI-3663.01), KIC 8827930 (orbital period 288 days, now KOI-3801.01), KIC 8813698 (orbital period 269 days, now KOI-1268.01), and KIC 9025971 (orbital period 141 days, now KOI-3680.01). The first three exhibited two transits in Q1–Q8 but were apparently not caught by eye. The fourth did not exhibit two transits, despite its shorter orbital period, due to missing data.
3. The Burke et al. (2014) sample, which is complete for Jupiter-sized planets that transit three times within Q1–Q8. KIC 8813698 (above) is now included as KOI-1268.01.
  4. The Q1–Q12 sample at the Exoplanet Archive. This sample now includes all four candidates we found on the TCE table: KIC 8813698 (now KOI-1268.01), KIC 8827930 (now KOI-3801.01), KIC 9025971 (now KOI-3680.01), and KIC 12735740 (now KOI-3663.01).
  5. The Q1–Q16 sample at the Exoplanet Archive, which is complete for Jupiter-sized planets that transit three times within Q1–Q16 ( $t_{\text{survey}} = 4$  yr).
  6. The catalog of long-period planets discovered by Planet Hunters (Wang et al. 2013; Schmitt et al. 2013). Of the 42 discoveries reported in Schmitt et al. (2013), three correspond to those we identified from the TCE table (KIC 8827930, 9025971, and 12735740). KIC 9413313 transits three times in Q1–Q16; its strong stellar variability may be the reason it did not appear on the Q1–Q16 candidate or TCE list. We add this candidate to our sample. Twenty-eight candidates fell below our planetary radius cut (KIC 2975770, 3326377, 3634051, 3732035, 4142847, 4472818, 48rd50, 4902202, 4947556, 5857656, 5871985, 5966810, 7826659, 8210018, 8636333, 9147029, 9166700, 9480535, 9886255, 10024862, 10360722, 10525077, 10850327, 11026582, 11253827.01, 11253827.02, 11392618, and 11716643). KIC 4760478b and the planet candidate orbiting 9663113 do not transit three times in Q1–Q16. KIC 8012732 does not have KIC stellar parameters available. For six others (KIC 3663173, 5437945.01, 5437945.02, 6878240, 9425139, and 9958387), when we fit the reported stellar parameters with the Takeda et al. (2007) models, we obtained stellar radii much smaller than those reported by (Wang et al. 2013), resulting in a planetary radius below our cut. Of the fourteen discoveries reported by Schmitt et al. (2013), four stars fall outside our stellar cuts (2437209, 5010054, 5522786, and 6805414), and eight have planets falling above or below our planetary radius cuts (5094412, 6436029, 9662267, 9704149, 10255705, 11152511, 11442793, and 12454613). KIC 5732155b does not transit three times in Q1–Q16. KIC 6372194b appears on the Q1–Q16 candidates list, but with a planetary radius below our cut due to the smaller planet-to-star radius ratio than that reported by Schmitt et al. (2013). We fit the light curves using the approach described in Section 3.1 and also found a small planet-to-radius ratio ( $R_p/R_\star = 0.085 \pm 0.003$ ), resulting in  $R_p < 8R_{\text{Earth}}$ . Therefore we do not include this candidate in our sample.
  7. The catalog by Huang et al. (2013) discovered using the HAT pipeline. Of the new candidates discovered by

Huang et al. (2013), only KIC 6805414 falls within our stellar and planetary cuts (KIC 5563300 falls below our planetary radius cut using the updated stellar parameters at the Exoplanet Archive). KIC 6805414 appears on the Q1–Q16 candidate list (KOI-5329.01) and may have been missed in Q1–Q12 by the *Kepler* pipeline because strong stellar variability occurs on a timescale similar to the transit duration. However, KIC 6805414 falls below our surface gravity cut using the updated parameters by Schmitt et al. (2013).

8. The catalogs of false positives (acquired from the NASA Exoplanet Archive) and eclipsing binaries (acquired from MAST). However, of the objects meeting our planetary radius cut, all had obvious secondary eclipses, even-odd eclipse depths, or centroid shifts. KOI-211.01 is on the list of *Kepler* eclipsing binaries, but since it may be a planet we keep it in our sample.
9. A catalog of long-period candidates created by an amateur astronomer (Person 2012). The Person (2012) catalog contains four candidates not in the Batalha et al. (2013) sample, three of which we identified above (KIC 12735740, KIC 8827930, and KIC 9025971) and one of which is the planet candidate identified by Huang et al. (2013), orbiting KIC 6805414.

Because these sources are consistent when compared across which quarters they probed, we have confidence that the giant planet candidates transiting three times are generally being picked up by the pipeline.

Our final list of long-period candidates is: KOI-209.01\*, KOI-211.01, KOI-351.01\*, KOI-372.01, KOI-398.01\*, KOI-458.01, KOI-682.01, KOI-806.01\*, KOI-806.02\*, KOI-815.01, KOI-918.0, KOI-1089.01\*, KOI-1193.01, KOI-1268.01, KOI-1353.01\*, KOI-1431.01, KOI-1439.01, KOI-1466.01, Kepler-419-b, KOI-1477.01, KOI-1483.01, KOI-1486.01\*, KOI-1552.01, KOI-1553.01, KOI-1587.01, KOI-2689.01, KOI-3680.01, KOI-3801.01, KOI-5071.01, KOI-5241.01, and KIC 9413313. Those with asterisks are in systems with multiple candidates.

Now we list a subset of candidates that appear on the candidate list, but that we exclude from our sample and describe our motivation for excluding them. The main reason why we end up excluding so many candidates is that the current *Kepler* catalog is very generous in including candidates that may be false positives, and the vetting of the current catalog is not uniform. We apply uniform vetting criteria to both the long-period planets and the hot Jupiters, making use of the data validation report.

We exclude the following candidates because they exhibit secondary eclipses inconsistent with planethood: KOI-193.01, KOI-856.01 (Ofir & Dreizler 2013), and KOI-3641.01. We exclude KOI-433.02 because it has large centroid-transit correlations (including an out-of-transit offset of 0.3 arcsec), as well as KOI-617.01 (0.7 arcsec), KOI-772.01 (0.2 arcsec), KOI-1645.01 (0.3 arcsec), KOI-1773.01 (0.3 arcsec) KOI-3678.01 (0.4 arcsec), KOI-3709.01 (0.7 arcsec), KOI-3717.01 (0.5 arcsec), and KOI-3726.01 (0.8 arcsec, and different odd/even eclipse depths). We also exclude KOI-620.02, KOI-625.01, KOI-1137.01, KOI-1242.01, KOI-1255.01 (which McQuillan et al. 2013 find also shows transit and rotation modulation on different stars), KOI-1684.01, KOI-1691.01, KOI-855.01, KOI-3660.01, KOI-4939.01, KOI-5312.01, KOI-5385.01, KOI-5488.01, and KOI-5792.01 due to strong centroid-flux correlations, which are both statistically significant and strikingly



evident by eye as a strong wind in flux versus position rain plot. We exclude KOI-1095.01 because it has large centroid (0.1 arcsec) and source (0.2 arcsec) offsets and a noticeable flux-centroid correlation; moreover, its radius is below 8 Earth radii in the Q1–Q12 and Q1–Q16 candidates list. KOI-366.01, KOI-377.02, KOI-777.01, KOI-1783.01, and KOI-1787.01 appeared on an earlier candidate list, but no longer meet our stellar and/or planetary criteria due to revised stellar properties. We do not include KOI-1288.01 because of revised stellar parameters (see Appendix F). KOI-1335.01 had a radius above 8 Earth radii in the Q1–Q8 candidate list, but in the Q1–Q12 candidate list, Q1–Q16 data validation report, and our own fits, it has a radius below 8 Earth radii, so we do not include it here. Candidate KOI-1496.01 is too small, and stars KOI-2659, KOI-5314, and KOI-5576 are too giant. (KOI-2569 was identified as a giant by McQuillan et al. 2013). KOI-375.01, KOI-422.01, and KOI-490.02 only transit twice in Q1–Q16; KOI-1096.01 only transit once. We do not include KOI-686.01, which we analyzed in detail in DJ12, because it was found by Díaz et al. (2014) to be a brown dwarf. KOI-771.01 and KOI-3787.01 have different odd-even eclipse depths and are likely false positives, so we do not include them here; KOI-5018.01 and KOI-5760.01 also have inconsistent depths. We also discard KOI-3320.01, which was originally on the eclipsing binaries list due to its large radius, because its position is highly offset from the KIC catalog, implying that the star is probably misidentified and therefore mischaracterized. Star KOI-3678 lacks KIC parameters. KOI-5409.01 only has one quarter of data and KOI-5682.01 only two. KOI-5446.01, KOI-5697.01, and KOI-5923.01 have low S/N (10, 9.4, and 7 respectively). KOI-5661.01 appears to be an instrumental artifact and KOI-5802.01 an artifact of stellar variability.

Several candidates that would cross the orbits of other candidates in their systems if they were highly eccentric also have very grazing transits that are difficult to fit (KOI-734.02, KOI-1258.03, KOI-1426.03) or are short period with many transits to extract (KOI-620.01); therefore we do not bother to fit them. We exclude KOI-1356.01 (which does not transit three times in Q1–Q12) because its true period is 787.45 days, which is far outside the completeness range we consider. (It is listed in the KIC with half this orbital period. If it had that orbital period, we would see a transit in Q16 but we do not.) We do not include KOI-872.01 (Kepler-46 b) because Nesvorný et al. (2012) constrain its eccentricity to be below 0.02 based on TTVs, or KOI-1574.01 (Kepler 87-b) because Ofir et al. (2014) constrain its eccentricity to be below 0.1 based on TTVs. We exclude KOI-3663.01 from our analysis (a.k.a., PH2 (Wang et al. 2013)) because Wang et al. (2013) measure its eccentricity to be  $e = 0.41^{+0.08}_{-0.29}$ ; thus it is not super-eccentric.

## APPENDIX B

### COMPUTING THE POSTERIOR OF THE MEAN NUMBER OF PLANETS BASED ON THE OBSERVED NUMBER OF PLANETS

The probability of observing  $N_{\text{pl}}$  from a Poisson distribution with mean  $\bar{N}_{\text{pl}}$  is:

$$\text{prob}(N_{\text{pl}}|\bar{N}_{\text{pl}}) = \frac{\bar{N}_{\text{pl}}^{N_{\text{pl}}}}{N_{\text{pl}}!} \exp[-\bar{N}_{\text{pl}}] \quad (\text{B1})$$

We wish to determine the posterior distribution for  $\bar{N}_{\text{pl}}$  and give the observed  $N_{\text{pl}}$ . Applying Bayes' theorem:

$$\begin{aligned} \text{prob}(\bar{N}_{\text{pl}}|N_{\text{pl}}) &= \text{prob}(N_{\text{pl}}|\bar{N}_{\text{pl}})\text{prob}(\bar{N}_{\text{pl}}) \\ &= \frac{\bar{N}_{\text{pl}}^{N_{\text{pl}}}}{N_{\text{pl}}!} \exp[-\bar{N}_{\text{pl}}]\text{prob}(\bar{N}_{\text{pl}}) \end{aligned} \quad (\text{B2})$$

where  $\text{prob}(\bar{N}_{\text{pl}})$  is the prior on  $\bar{N}_{\text{pl}}$ . For a uniform prior on  $\text{prob}(\bar{N}_{\text{pl}})$ , the median of the posterior,  $\text{Med}(\bar{N}_{\text{pl}})$  is the solution to the equation:

$$0.5 = \frac{\int_{\text{Med}(\bar{N}_{\text{pl}})}^{\infty} \frac{\bar{N}_{\text{pl}}^{N_{\text{pl}}}}{N_{\text{pl}}!} \exp[-\bar{N}_{\text{pl}}]\text{prob}(\bar{N}_{\text{pl}})d\bar{N}_{\text{pl}}}{\int_0^{\infty} \frac{\bar{N}_{\text{pl}}^{N_{\text{pl}}}}{N_{\text{pl}}!} \exp[-\bar{N}_{\text{pl}}]\text{prob}(\bar{N}_{\text{pl}})d\bar{N}_{\text{pl}}} \quad (\text{B3})$$

For a uniform prior  $\text{prob}(\bar{N}_{\text{pl}}) \propto 1$ ,

$$0.5 = \frac{\int_{\text{Med}(\bar{N}_{\text{pl}})}^{\infty} \bar{N}_{\text{pl}}^{N_{\text{pl}}} \exp[-\bar{N}_{\text{pl}}]d\bar{N}_{\text{pl}}}{\int_0^{\infty} \bar{N}_{\text{pl}}^{N_{\text{pl}}} \exp[-\bar{N}_{\text{pl}}]d\bar{N}_{\text{pl}}} = \frac{\Gamma[N_{\text{pl}} + 1, \text{Med}(\bar{N}_{\text{pl}})]}{\Gamma[N_{\text{pl}} + 1]} \quad (\text{B4})$$

where  $\Gamma$  is the gamma function.

For a Jeffrey's prior (appropriate when the scale of the parameter is unknown), which for a Poisson distribution is  $\text{prob}(\bar{N}_{\text{pl}}) \propto (\bar{N}_{\text{pl}})^{-1/2}$  (e.g., Bernardo & Smith 2000; Farr et al. 2013; Tremmel et al. 2013),

$$\begin{aligned} 0.5 &= \frac{\int_{\text{Med}(\bar{N}_{\text{pl}})}^{\infty} \bar{N}_{\text{pl}}^{N_{\text{pl}}-0.5} \exp[-\bar{N}_{\text{pl}}]d\bar{N}_{\text{pl}}}{\int_0^{\infty} \bar{N}_{\text{pl}}^{N_{\text{pl}}-0.5} \exp[-\bar{N}_{\text{pl}}]d\bar{N}_{\text{pl}}} \\ &= \frac{\Gamma[N_{\text{pl}} + 0.5, \text{Med}(\bar{N}_{\text{pl}})]}{\Gamma[N_{\text{pl}} + 0.5]} \end{aligned} \quad (\text{B5})$$

See Section 3.4 of Farr et al. (2013) for a detailed discussion of the appropriate prior for Poisson counting statistics. The 68.3% confidence interval is calculated by equating the ratios in Equation (B4) and (B5) to 0.1585 and 0.8415 (i.e.,  $0.5 \pm 0.683/2$ ). These posteriors have medians slightly larger than the counted numbers because of the skewed shape of a Poisson distribution at small values of the mean ( $\bar{N}_{\text{pl}} < 10$ ). It is more probable that we are observing fewer planets than the true mean number than vice versa. As an extreme example, if the mean number of planets per sample is greater than zero, there is some possibility that our sample will happen to contain zero.

## APPENDIX C

### SHORT PERIOD PLANET SAMPLE

Here we list the planets from non-*Kepler* surveys that make up the sample in Table 1:

1. Moderately eccentric Jupiters ( $N_{\text{mod},0}$ ) with  $2.8 < P_{\text{final}} < 5$  days: CoRoT-16-b, HAT-P-2-b, HAT-P-21-b, HAT-P-31-b, HAT-P-34-b, and XO-3-b.
2. Moderately eccentric Jupiters ( $N_{\text{mod},0}$ ) with  $5 < P_{\text{final}} < 10$  days: CoRoT-10-b, HAT-P-17-b, WASP-8-b, HD-185269-b, HD-118203-b, HD-162020-b, HD-108147-b
3. Non-*Kepler* Jupiters with  $2.8 < P < 5$  days ( $N_{P=P_{\text{final},0}}$ ): CoRoT-12-b, CoRoT-13-b, CoRoT-17-b, CoRoT-19-b,

- CoRoT-23-b, CoRoT-5-b, HAT-P-1-b, HAT-P-12-b, HAT-P-13-b, HAT-P-19-b, HAT-P-20-b, HAT-P-21-b, HAT-P-22-b, HAT-P-24-b, HAT-P-25-b, HAT-P-27-b, HAT-P-28-b, HAT-P-30-b, HAT-P-33-b, HAT-P-35-b, HAT-P-38-b, HAT-P-39-b, HAT-P-4-b, HAT-P-8-b, HATS-1-b, KELT-2 A-b, OGLE-TR-211-b, WASP-10-b, WASP-11-b, WASP-15-b, WASP-16-b, WASP-21-b, WASP-22-b, WASP-25-b, WASP-29-b, WASP-31-b, WASP-34-b, WASP-35-b, WASP-37-b, WASP-41-b, WASP-42-b, WASP-45-b, WASP-47-b, WASP-55-b, WASP-6-b, WASP-61-b, WASP-62-b, WASP-63-b, WASP-67-b, WASP-7-b, WTS-1-b, XO-3-b, tau-boo-b, 51 Peg-b, HD-102195-b, HD-149143-b, HD-149026-b, HD-179949-b, HD-187123-b, HD-209458-b, HD-2638-b, HD-330075-b, HD-63454-b, HD-75289-b, HD-83443-b, HD-88133-b, upsilon And-b, P-201-b, and BD-103166-b. We do not include the following planets because their eccentricities are poorly constrained (and therefore may either be above or below  $e = 0.2$ ): CoRoT-11-b, HAT-P-3-b, HAT-P-9-b, OGLE-TR-10-b, OGLE-TR-111-b, OGLE-TR-182-b, TrES-1-b, TrES-4-b, WASP-13-b, WASP-39-b, WASP-58-b, WASP-60-b, XO-1-b, XO-4-b, and XO-5-b. The following planets have  $e$  fixed at zero in the OD, but are constrained in the literature to be below  $e < 0.2$  and therefore we include them in the sample: HAT-P-1-b, HAT-P-4-b, HAT-P-8-b, HAT-P-12-b, HAT-P-27-b, HAT-P-39-b, OGLE-TR-211-b, KELT-2-Ab, WASP-7, WASP-11-b, WASP-15-b, WASP-21-b, WASP-25-b, WASP-31-b, WASP-35-b, WASP-37-b, WASP-41-b, WASP-42-b, WASP-47-b, WASP-61-b, WASP-62-b, WASP-63-b, and WASP-67-b. Finally, we constrain the eccentricities of CoRoT-13-b (Cabrera et al. 2010), CoRoT-17-b (Csizmadia et al. 2011), and WASP-16-b (Lister et al. 2009) to be  $e < 0.2$  by fitting the radial-velocity data, following Dawson & Fabrycky (2010) except using the MCMC algorithm (instead of the Levenberg-Marquardt algorithm) to obtain posteriors for the eccentricities.
4. *Kepler* Jupiters with  $2.8 < P < 5$  days ( $N_{P=P_{\text{final}}}$ ): KOI-10.01, KOI-17.01, KOI-18.01, KOI-20.01, KOI-97.01, KOI-127.01, KOI-128.01, KOI-135.01, KOI-186.01, KOI-188.01, KOI-195.01, KOI-199.01, KOI-201.01, KOI-204.01, KOI-217.01, KOI-421.01, KOI-611.01, KOI-760.01, KOI-767.01, KOI-830.01, KOI-838.01, KOI-908.01, KOI-913.01, and KOI-1074.01. We do not include two known false positives (KOI-208.01 and KOI-895.01, Demory & Seager 2011). We do not include KOI-931.01 because it exhibits a very large (1.7 arcsecond) centroid offset during transit, indicating a likely blend/false positive. For the same reason we do not include KOI-214.01 or KOI-813.01. KOI-554.01 was found to be a brown dwarf by Santerne et al. (2012).
  5. Non-*Kepler* Jupiters with  $5 < P < 10$  days ( $N_{P=P_{\text{final},0}}$ ): CoRoT-16-b, CoRoT-4-b, CoRoT-6-b, HAT-P-18-b, HAT-P-2-b, HAT-P-29-b, HAT-P-31-b, HAT-P-34-b, WASP-38-b, WASP-59-b, WASP-8-b, HD-185269-b, HD-118203-b, HIP 14810-b, HD-68988-b, HD-162020-b, HD-217107-b, HD-109749-b.
  6. *Kepler* Jupiters with  $5 < P < 10$  days ( $N_{P=P_{\text{final}}}$ ): KOI-22.01, KOI-98.01, KOI-131.01, KOI-200.01, KOI-206.01, KOI-428.01, KOI-466.01, KOI-680.01, KOI-728.01, KOI-774.01, KOI-889.01, KOI-890.01, KOI-929.01, KOI-1391.01, KOI-1456.01, and KOI-1465.01. We do not include known false positives KOI-425.01 or KOI-607.01 (Santerne et al. 2012 and see also Madhusudhan et al. 2012).

We exclude KOI-3627.02 because it appears to be a false positive. We do not include KOI-1066.01 because it exhibits a very large (3.1 arcsecond) centroid offset during transit, indicating a likely blend/false positive. For the same reason we don't include KOI-1457.01. We exclude KOI-3721.01, KOI-3767.01, and KOI-3771.01 because they were only observed starting in Q10 and have extreme stellar variability, making it highly unlikely that their super-eccentric prototypes would have been detected. We also exclude KOI-3689.01 because it was only observed during Q10.

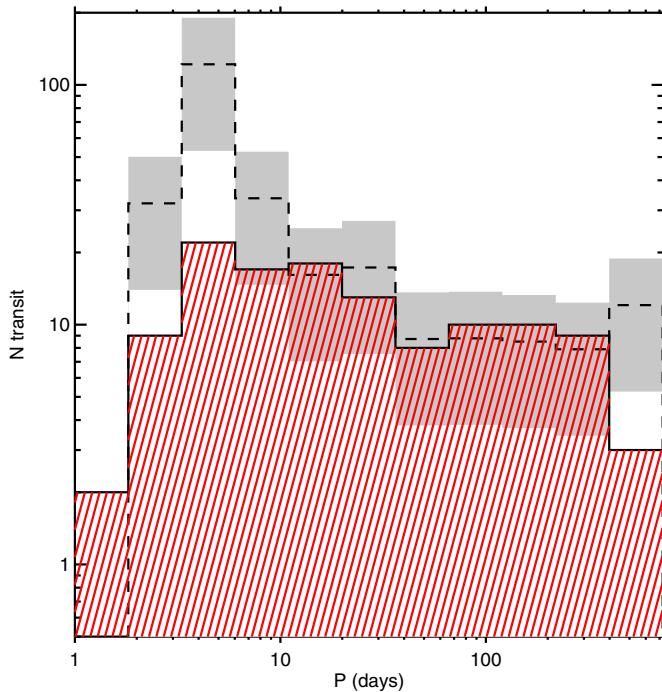
These numbers differ from those in S12 for several reasons:

1. S12 used the same definition that we use  $N_{P=P_{\text{final}}}$  for the *Kepler* sample but, for the calibration sample, they used the total number of planets observed along the  $P_{\text{final}}$  track. However, we wish to treat both samples the same and therefore use the same class of object for both:  $N_{P=P_{\text{final}}}$  for the *Kepler* sample and  $N_{P=P_{\text{final},0}}$  for the calibration sample.
2. We combine non-*Kepler* planets detected by transit surveys and those detected radial-velocity surveys in order to enhance our sample size and get a better estimate for Equation (1), rather than treating the two samples separately. This is justifiable because even though the total number of planets along a given angular momentum track has different selection biases for transit versus radial-velocity, the *fraction* of moderately eccentric Jupiters along a given track should be consistent in the two samples (as S12 found to be the case). The non-*Kepler* transit surveys are not better suited than the radial-velocity surveys for a comparison to the *Kepler* stars, except for the transit probability. As noted by S12, the transit probability is constant along an angular momentum track. Therefore the ratio of planets along different portions of a given track should not differ between radial velocity versus transit samples.
3. We impose cuts on stellar effective temperature and surface gravity to restrict our sample to the type of stars with reliable stellar density estimates.
4. Additional non-*Kepler* planets have been discovered and characterized since S12.
5. Some of the *Kepler* hot Jupiter candidates had their status or parameters changed. Of the 46 hot Jupiters in the Borucki et al. (2011) sample used by S12, 18 are no longer included: one had its radius decreased below 8 Earth radii (KOI-214.01), 11 were designated a false positive (KOI-194.01, KOI-552.01, KOI-609.01, KOI-822.01, KOI-840.01, KOI-876.01, KOI-1003.01, KOI-1152.01, KOI-1177.01, KOI-1382.01, KOI-1543.01), three were declared a false positive in the literature (KOI-208.01, KOI-425.01, KOI-895.01), and one (KOI-931.01) was declared a likely false positive above. KOI-410.01 has a radius of 45 Earth radii and KOI-684 of 30 Earth radii, above our upper limit of 22 Earth radii.

## APPENDIX D

### COMPLETENESS OF THE *KEPLER* SAMPLE

We compare the *Kepler* and radial-velocity samples, following Dawson & Murray-Clay (2013) Section 4. We plot (Figure 8) the actual observed number of *Kepler* giant planets and compare the number expected from the radial-velocity sample (i.e., those



**Figure 8.** Red striped: number of transiting giant planets detected by *Kepler*, using the stellar and planetary cuts described in this work and removing the false positives mentioned in Appendices A and C. Black dashed: expected number based on the RV-discovered (i.e., excluding planets discovered by transit surveys) sample. The gray error bars are from uncertainties in the normalization, not the Poisson uncertainties of each individual bin.

planets plotted in Figure 1),

$$N_{RV,trans} = C_{norm} N_{RV} \text{prob}_{trans},$$

where  $N_{RV}$  is the observed number of RV planets per bin and  $\text{prob}_{trans}(P)$  is the transit probability. We normalize the radial-velocity sample using the absolute occurrence rates of hot Jupiters, which we expect to be complete, and our analytical completeness estimate from Section 2.1. We use the RV hot Jupiter occurrence rate from Wright et al. (2012) ( $f_{HJ,RV} = 1.2 \pm 0.38\%$ ) and the *Kepler* hot Jupiter occurrence rate from Fressin et al. (2013) ( $f_{HJ,Kepler} = 0.43 \pm 0.05\%$ ). The normalization constant for the RV sample is:

$$C_{norm} = \frac{f_{HJ,RV}}{f_{HJ,Kepler}} \frac{\sum_{P=0}^{10 \text{ days}} N_{trans,Kepler}(P)/\text{prob}_{trans}(P)}{\sum_{P=0}^{10 \text{ days}} N_{RV}(P)} C_{comp}$$

where  $C_{comp}$  is defined in Section 2.1. We note that the comparison here is intended only as a consistency check; a detailed comparison of the *Kepler* and radial-velocity giant planet period distributions would require careful thought about the mapping between mass and radius, the observational biases in both samples (apart from the transit probability), and other considerations.

At the orbital periods considered here, ( $P > 30$ ) days, the distributions appear consistent except in the period bin spanning 398–724 days. There are three transiting Jupiters in this bin, whereas from the RV sample we would expect between 5 and 19 (where the range represents uncertainty in the normalization, not Poisson uncertainties). However, given Poisson uncertainties (for which an observed 3 represents a true number 2 to 5 at one sigma; see Appendix B), this is consistent with the lower end of the normalization range. Moreover, this period corresponds to the dramatic uptick in planet frequency observed in the radial

velocity sample (Cumming et al. 2008), and the intrinsic period distribution in the *Kepler* sample may be different than in the radial-velocity sample. We are developing our own pipeline into which we can inject transits to get a better estimate of completeness for a follow-up to this study.

## APPENDIX E

### AVOIDING PROBLEMS DUE TO INCORRECT STELLAR PARAMETERS

Characterizing the entire planetary eccentricity distribution from transit light curve parameters can be complicated by systematic errors and uncertainties in the stellar parameters (e.g., Moorhead et al. 2011; Kane et al. 2012; Plavchan et al. 2014). Instead, we simply aim to determine whether or not there are light curves for which  $\rho_{circ}$  is physically unlikely. For example, a planet transiting at periape with  $e = 0.95$  would have  $\rho_{circ} = 244\rho_*$ . Such a high density would be astrophysically implausible based on our knowledge of stellar evolution.

We consider three potential problems caused by incorrect stellar parameters. First, we could mistakenly identify a planet as being highly eccentric even if it had  $\rho_{circ} \sim \rho_\odot$  if we were to underestimate  $\rho_*$  as being very low. Second, we could miss an eccentric planet if we thought its host star had  $\rho_* \sim \rho_\odot$  but the true stellar density were much smaller. However, we avoid both of these problems by restricting our samples to exclude giants. Regarding the first problem, all of the  $\rho_*$  we derive for *Kepler* hosts in Section 3 are of order 1. Regarding the second problem, by excluding giants from the well-characterized calibration sample of stars with both transits and RV measurements and from the *Kepler* hot Jupiter sample (Figure 1), we did not make any predictions for super-eccentric planets orbiting giants, and therefore cannot miss any.

The third potential problem results from using the number of hot Jupiters in the *Kepler* sample,  $N_{P=P_{final}}$ , as an input for predicting  $N_{sup}$  (Equation (2)). If a large fraction of the *Kepler* hot Jupiters orbits stars that are secretly giants but slipped past our stellar parameter cuts, then we might overpredict the expected number of super-eccentric proto-hot Jupiters. As shown by Mann et al. (2012) and Dressing & Charbonneau (2013), some of the stars classified as M or K dwarfs might be giants. However, Mann et al. (2012) find that imposing a cut of  $\log g > 4$ , as we do, helps avoid this misclassification. With the cut imposed, 97% of cool stars dimmer than  $Kp = 14$  are dwarves (Mann et al. 2012). Among our sample of *Kepler* hot Jupiters, all the stars with  $T_{eff} < 5714$  K have  $Kp > 14$ , so it is very unlikely our sample harbors many giants masquerading as M or K dwarves.

Recently, K14 presented several caveats in measuring planet eccentricities using posteriors of  $\rho_{circ}$  and  $\rho_*$ . The measurement of  $\rho_{circ}$  from the light curve can be affected by blends, spots, transit timing variations (TTVs), transit duration variations (TDVs), and the planet's mass. K14 shows that a blend factor  $B$  reduces the observed  $\rho_{circ}$  by a factor of  $B^{-3/4}$ . In order to fail to identify a candidate as being super-eccentric, the blend would have to be of order 10. However, this would mean that the observed transit depth has been reduced by a factor of order 3 and that the true transit is not of a planet but of an eclipsing binary star. Therefore, since we are searching for super-eccentric planets not stars, this is not the type of object we are looking for. Conversely, K14 demonstrates that spots cause  $\rho_{circ}$  to be overestimated; therefore this effect would not prevent us from identifying super-eccentric planets. We avoid problems caused



by TTVs and TDVs by fitting a separate transit time and impact parameter for each transit when justified (e.g., for Kepler-419-b). Finally, K14 shows that the planet’s mass only affects  $\rho_{\text{circ}}$  by  $1 + M_p/M_*$ , which is small compared to our uncertainties in  $\rho_{\text{circ}}$ .

A final caveat by K14 is the limitation of the small-angle approximation used to link the eccentricity  $e$  to  $\rho_{\text{circ}}$  and  $\rho_*$  (e.g., as in DJ12). Condition B (Equation (B11)) of K14 is that

$$\sin x \approx x \quad (\text{E1})$$

for which

$$x = \frac{q_c^2}{\sqrt{1-e^2}} \sqrt{\frac{(1 \pm p)^2 - b^2}{(a/R_*)^2 q_c^2 - b^2}} \quad (\text{E2})$$

for which (K14, Equation (B4)):

$$q_c = \frac{1 - e^2}{1 + e \sin \omega}. \quad (\text{E3})$$

After rewriting Equation (E3) in terms of our variables  $g$  and  $a_{\text{final}}$

$$q_c = \frac{\sqrt{a_{\text{final}}/a}}{g}, \quad (\text{E4})$$

we can rewrite Equation (E2) as:

$$x = \frac{\left(\frac{\sqrt{a_{\text{final}}/a}}{g}\right)^2}{\sqrt{a_{\text{final}}/a}} \sqrt{\frac{(1 \pm p)^2 - b^2}{(a/R_*)^2 \left(\frac{\sqrt{a_{\text{final}}/a}}{g}\right)^2 - b^2}} \quad (\text{E5})$$

$$x = \frac{\sqrt{a_{\text{final}}/a}}{g^2} \sqrt{\frac{(1 \pm p)^2 - b^2}{\left(\frac{aa_{\text{final}}}{R_*^2 g^2}\right) - b^2}}.$$

In order for  $\sin x \approx x$ , we need  $x \ll \sqrt{6}$  (Equation (B13) of K14). For  $\sqrt{a_{\text{final}}/a} < 1$  and  $g > 1$ , we are only in danger of violating this condition when the denominator is small, i.e.,  $(\sqrt{aa_{\text{final}}/R_*g}) \sim b$ . In our search for super-eccentric Jupiters, we consider  $a_{\text{final}} > 8R_*$  and  $a > 50R_*$ . Therefore, in order for the LHS to be comparable to an impact parameter (which ranges between 0 and 1), we would need  $g > 19$ . We do not have any planets like this in our sample; the biggest  $g$  in our sample is of order 2.5.

## APPENDIX F

### EXCEPTIONAL CANDIDATE TREATMENTS

As part of our spectroscopic survey, we measured spectroscopic parameters using SpecMatch for KOI-211 and Kepler-419, which we use instead of the KIC parameters. Everett et al. (2013) conducted a spectroscopic survey of faint KOI. When the Huber et al. (2014) parameters are derived from photometry and the Everett et al. (2013) parameters are derived from spectroscopy, we use the Everett et al. (2013) parameters instead. This applies to KOI-209, KOI-398, KOI-918, KOI-1089, and KOI-1268. They found that KOI-1288 has  $\log g < 4$ ; since this places the star outside the range of stellar parameters we consider, we do not include it in our sample. We also checked if any of the *Kepler* hot Jupiters had their stellar parameters revised outside our range, but none did.

The Takeda et al. (2007) evolution models only include stars with  $M_* > 0.7 M_\odot$ . For a subset of low-mass host stars—KOI-1466, KOI-1477, and KOI-1552—we use the Dartmouth

stellar evolution models (Dotter et al. 2008) instead. We use models with the default helium core mass and solar  $\alpha$  abundance. We interpolate and resample to maintain a uniform prior in stellar age, mass, and metallicity.

Based on the V-shapes of their light curves, several of the candidates are probably not planets but grazing eclipsing binaries (KOI-1193.01, KOI-1483.01, KOI-1587.01, KOI-5071.01, and KOI-5241.01). Of the two of these that were candidates in the Batalha et al. (2013) catalog, both are noted as V-shaped in that catalog. Although we include them here, we only use the fraction of the posterior for which  $R_p < 22R_\oplus$ . Moreover, we use this fraction to weight the contribution to the total number of super-eccentric Jupiters in Section 3.2.

## APPENDIX G

### ASSUMPTIONS THAT CANNOT EXPLAIN A LOWER THAN EXPECTED NUMBER OF SUPER-ECCENTRIC PROTO-HOT JUPITERS

These assumptions cannot explain observing fewer than expected super-eccentric proto-hot Jupiters, either because a violation would result in *more* super-eccentric progenitors (1) or because they are unlikely to be violated (2)–(5).

1. *The evolution of the planet’s radius due to tidal inflation is negligible, and no planets are disrupted by tides.* However, if the planet’s radius were to expand over the course of HEM due to tidal inflation, then the tidal dissipation rate would be even lower during the earlier stages of HEM, causing planets to spend more time at high eccentricities. Therefore, this effect could only increase the expected number of super-eccentric proto-hot Jupiters. The prevalence of tidal disruption does not affect the S12 prediction, because the prediction is based on the survivors. Depending on the timescale of tidal disruption, we may observe additional doomed proto-hot Jupiters that will not survive their HEM.
2. *Angular momentum is not exchanged between the planet and star.* If planets were to typically transfer angular momentum to stars, we would expect more super-eccentric hot Jupiters than predicted and vice versa. However, Penev et al. (2012) argue that stellar tidal dissipation is likely unimportant, because if it were, most hot Jupiters would be subsumed by their stars on short timescales. We note that although a star can add or remove angular momentum from the planet’s orbit as the star rapidly expands on the giant branch (e.g., Santerne et al. 2012), we strictly restrict our samples to main-sequence stars so we can ignore this effect. We note that the Sun’s spin angular momentum ranges from about 10% ( $P_{\text{final}} = 3$  days) to 6% ( $P_{\text{final}} = 10$  days) of the orbital angular momentum of a proto-hot Jupiter.
3. *The planet’s orbital angular momentum and spin angular momentum are not exchanged.* We neglect this effect because we assume that the planet maintains a pseudo-synchronous spin throughout its evolution. If the ratio of the planet’s orbital angular momentum to its spin angular momentum is large, the planet’s spin quickly (compared to the circularization timescale) evolves to this pseudo-synchronous state, in which the planet’s spin rate is similar to the orbital frequency at periape. We expect that the ratio of orbital to spin angular momentum is indeed typically large because the planet’s distance from the star is very large compared to the planetary radius.
4. *Moderately eccentric calibration Jupiters ( $N_{\text{mod},0}$ ) truly have  $e > 0.2$ ; they are not low-eccentricity planets that*

appear eccentric due to eccentricity bias. Eccentricity bias occurs when noise masquerades as eccentricity. Because the eccentricity cannot be negative, it is biased toward higher values. If one decomposes the RV signal caused by an eccentric planet into sinusoidal harmonics of the planet's orbital frequency, one finds that the signal due to eccentricity is primarily embedded in the second harmonic and has an amplitude of  $eK$ , where  $K$  is the RV amplitude (e.g., Anglada-Escudé et al. 2010). Eccentricity bias is primarily a concern when  $eK$  is near the noise level (i.e., for low-mass and/or long-period planets with small  $K$ ). In contrast, Jupiter-mass planets on short-period orbits have large  $K$ . For an RV precision of a few m/s and a typical hot Jupiter  $K \sim 100$  m/s, a signal of amplitude  $e \times K = 0.2 \times 100$  m/s = 20 m/s is well above the noise level. Moreover, an even tighter constraint on the planet's eccentricity is possible through a joint fit to the RVs and transit light curve, as performed for a number of planets in the calibration sample. Therefore we expect that the calibration sample of moderately eccentric Jupiters (which have orbital periods ranging from 3 to 15 days) truly has  $e > 0.2$ .

5. *Only a small fraction of Kepler hot Jupiters are false positives.* The expected number of proto-hot Jupiters is proportional to the true rate of *Kepler* hot Jupiters  $\bar{N}_{P=P_{\text{final}}}$  (Equation (2)). For example, if half the *Kepler* hot Jupiters were false positives, the predicted number of proto-hot Jupiters should be cut in half. Morton & Johnson (2011b) and Desert et al. (2012) find low false-positive rates for *Kepler* candidates (<10%). Santerne et al. (2012), Colón et al. (2012), and Fressin et al. (2013) find higher false-positive rates. However, Santerne et al. (2012) focused on a population with high a priori false-positive probabilities because of their V-shaped light curves. Moreover, the false positive discoveries by Colón et al. (2012) were for planets with  $P < 3$  days, which Colón et al. (2012) suggested can be expected from the period distribution of binaries, and we do not include planets with  $P < 3$  days in our sample here. We did not include the known false-positives in our computations in Section 2. Finally, the false positive rate derived by Fressin et al. (2013) is somewhat larger (18%). However, this rate was based on the Batalha et al. (2013) sample and, since then, roughly 25% of hot Jupiters have been removed from the *Kepler* sample and marked as false positives (Burke et al. 2014), so we expect that the false positive rate of the Burke et al. (2014) sample that we use is significantly lower.

Finally, S12 assumed that tidal evolution occurs according to the constant tidal time lag approximation (Hut 1981; Eggleton et al. 1998; Socrates & Katz 2012; Socrates et al. 2012a). This assumption controls the ratio  $r(e_{\text{max}})$  (Equation (3)) of high-eccentricity proto-hot Jupiters to moderate-eccentricity hot Jupiters along a given angular momentum track. The constant tidal time lag approximation is conventional but may be violated: if the dissipation rate were larger for highly eccentric Jupiters along a given angular momentum track than for moderately eccentric Jupiters than predicted by the constant time-lag model, we would expect fewer super-eccentric proto-hot Jupiters than predicted or vice versa. Hansen (2012) makes the argument that tidal dissipation is stronger at longer orbital periods because the forcing period is longer. However, in the highly eccentric regime we consider here, the forcing period is  $(1/2)P_{\text{final}}$  (e.g., Wu & Murray 2003, Section 3.3). Therefore the forcing period is constant and we expect that if equilibrium tides dominate, the

ratio of high-eccentricity to moderate-eccentricity Jupiters used here is not too low.

Dynamical tides, in which dissipation occurs through surface gravity waves (e.g., Zahn 1975), may be important in a proto-hot Jupiter's tidal evolution. Beaugé & Nesvorný (2012) argue that dynamical tides act at high eccentricities and equilibrium tides at low eccentricities; they added an empirical correction factor to the constant time-lag model so that, at large eccentricities, it matches the numerical results of the dynamical tide model computed by Ivanov & Papaloizou (2011). The empirical correction factor is proportional to  $10^{200qe^2}$ , where  $q$  is the periape distance (Beaugé & Nesvorný 2012). Along a given angular momentum track,  $q = a_{\text{final}}/(1+e)$ , so the tidal dissipation timescale [proportional to  $10^{200a_{\text{final}}e^2/(1+e)}$ ] is longer for larger eccentricities. If this correction factor applies, the contribution of dynamical tides would increase the expected number of super-eccentric proto-hot Jupiters.

The effect of tides on orbital evolution remains uncertain and is a topic of ongoing research. It remains unclear whether the true typical tidal evolution would result in more or in fewer super-eccentric Jupiters. Our results should be revisited as new theories and formulations of tidal evolution are developed.

## REFERENCES

- Albrecht, S., Winn, J. N., Johnson, J. A., et al. 2012, *ApJ*, **757**, 18
- Alibert, Y., Mordasini, C., Benz, W., & Winisdoerffer, C. 2005, *A&A*, **434**, 343
- Anglada-Escudé, G., López-Morales, M., & Chambers, J. E. 2010, *ApJ*, **709**, 168
- Barnes, J. W. 2007, *PASP*, **119**, 986
- Barnes, R., & Raymond, S. N. 2004, *ApJ*, **617**, 569
- Batalha, N. M., Rowe, J. F., Bryson, S. T., et al. 2013a, *ApJS*, **204**, 24
- Batalha, N. M., Rowe, J. F., Bryson, S. T., et al. 2013b, *ApJS*, **204**, 24
- Beaugé, C., & Nesvorný, D. 2012, *ApJ*, **751**, 119
- Bernardo, J., & Smith, A. 2000, *Bayesian Theory* (New York: Wiley)
- Boley, A. C., Payne, M. J., & Ford, E. B. 2012, *ApJ*, **754**, 57
- Borucki, W. J., Koch, D. G., Basri, G., et al. 2011, *ApJ*, **736**, 19
- Bromley, B. C., & Kenyon, S. J. 2011, *ApJ*, **735**, 29
- Brown, T. M., Latham, D. W., Everett, M. E., & Esquerdo, G. A. 2011, *AJ*, **142**, 112
- Burke, C. J., Bryson, S., Mullally, F., et al. 2014, *ApJS*, **210**, 19
- Cabrera, J., Bruntt, H., Ollivier, M., et al. 2010, *A&A*, **522**, A110
- Carter, J. A., & Winn, J. N. 2009, *ApJ*, **704**, 51
- Chatterjee, S., Ford, E. B., Matsumura, S., & Rasio, F. A. 2008, *ApJ*, **686**, 580
- Colón, K. D., Ford, E. B., & Morehead, R. C. 2012, *MNRAS*, **426**, 342
- Csizmadia, S., Moutou, C., Deleuil, M., et al. 2011, *A&A*, **531**, A41
- Cumming, A., Butler, R. P., Marcy, G. W., et al. 2008, *PASP*, **120**, 531
- Dawson, R. I., & Fabrycky, D. C. 2010, *ApJ*, **722**, 937
- Dawson, R. I., & Johnson, J. A. 2012, *ApJ*, **756**, 122
- Dawson, R. I., Johnson, J. A., Fabrycky, D. C., et al. 2014, *ApJ*, **791**, 89
- Dawson, R. I., Johnson, J. A., Morton, T. D., et al. 2012, *ApJ*, **761**, 163
- Dawson, R. I., & Murray-Clay, R. A. 2013, *ApJL*, **767**, L24
- Demory, B.-O., & Seager, S. 2011, *ApJS*, **197**, 12
- Desert, J.-M., Charbonneau, D., Fressin, F., & Torres, G. 2012, *BAAS*, **219**, 414.02
- Díaz, R. F., Montagnier, G., Leconte, J., et al. 2014, arXiv:1410.5248
- Dong, S., Katz, B., & Socrates, A. 2013, *ApJL*, **763**, L2
- Dong, S., Katz, B., & Socrates, A. 2014, *ApJL*, **781**, L5
- Dotter, A., Chaboyer, B., Jevremović, D., et al. 2008, *ApJS*, **178**, 89
- Dressing, C. D., & Charbonneau, D. 2013, *ApJ*, **767**, 95
- Dunhill, A. C., Alexander, R. D., & Armitage, P. J. 2013, *MNRAS*, **428**, 3072
- Eggleton, P. P., Kiseleva, L. G., & Hut, P. 1998, *ApJ*, **499**, 853
- Everett, M. E., Howell, S. B., Silva, D. R., & Szkody, P. 2013, *ApJ*, **771**, 107
- Fabrycky, D., & Tremaine, S. 2007, *ApJ*, **669**, 1298
- Fabrycky, D. C., & Winn, J. N. 2009, *ApJ*, **696**, 1230
- Farr, W. M., Gair, J. R., Mandel, I., & Cutler, C. 2013, arXiv:1302.5341
- Ford, E. B., Quinn, S. N., & Veras, D. 2008, *ApJ*, **678**, 1407
- Ford, E. B., & Rasio, F. A. 2006, *ApJL*, **638**, L45
- Ford, E. B., & Rasio, F. A. 2008, *ApJ*, **686**, 621
- Fressin, F., Torres, G., Charbonneau, D., et al. 2013, *ApJ*, **766**, 81
- Gaudi, B. S., Seager, S., & Mallen-Ornelas, G. 2005, *ApJ*, **623**, 472

- Gazak, J. Z., Johnson, J. A., Tonry, J., et al. 2012, *AdAst*, 2012, 697967
- Girardi, L., Groenewegen, M. A. T., Hatziminaoglou, E., & da Costa, L. 2005, *A&A*, **436**, 895
- Goldreich, P., Lithwick, Y., & Sari, R. 2004, *ARA&A*, **42**, 549
- Goldreich, P., & Sari, R. 2003, *ApJ*, **585**, 1024
- Goldreich, P., & Tremaine, S. 1980, *ApJ*, **241**, 425
- Guillochon, J., Ramirez-Ruiz, E., & Lin, D. 2011, *ApJ*, **732**, 74
- Hansen, B. M. S. 2010, *ApJ*, **723**, 285
- Hansen, B. M. S. 2012, *ApJ*, **757**, 6
- Hébrard, G., Almenara, J.-M., Santerne, A., et al. 2013, *A&A*, **554**, A114
- Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2012, *ApJS*, **201**, 15
- Huang, X., Bakos, G. A., & Hartman, J. D. 2013, *MNRAS*, **429**, 2001
- Huber, D., Silva Aguirre, V., Matthews, J. M., et al. 2014, *ApJS*, **211**, 2
- Hut, P. 1981, *A&A*, **99**, 126
- Ida, S., & Lin, D. N. C. 2008, *ApJ*, **673**, 487
- Ida, S., Lin, D. N. C., & Nagasawa, M. 2013, *ApJ*, **775**, 42
- Ivanov, P. B., & Papaloizou, J. C. B. 2011, *CeMDA*, **111**, 51
- Jones, H. R. A., Butler, R. P., Tinney, C. G., et al. 2003, *MNRAS*, **341**, 948
- Jurić, M., & Tremaine, S. 2008, *ApJ*, **686**, 603
- Kane, S. R., Ciardi, D. R., Gelino, D. M., & von Braun, K. 2012, *MNRAS*, **425**, 177
- Kipping, D. M. 2008, *MNRAS*, **389**, 1383
- Kipping, D. M. 2010a, *MNRAS*, **408**, 1758
- Kipping, D. M. 2010b, *MNRAS*, **407**, 301
- Kipping, D. M. 2014, *MNRAS*, **440**, 2164
- Lee, E. J., Chiang, E., & Ormel, C. W. 2014, *ApJ*, **797**, 95
- Lissauer, J. J., Ragozzine, D., Fabrycky, D. C., et al. 2011, *ApJS*, **197**, 8
- Lister, T. A., Anderson, D. R., Gillon, M., et al. 2009, *ApJ*, **703**, 752
- Lithwick, Y., & Naoz, S. 2011, *ApJ*, **742**, 94
- Madhusudhan, N., Lee, K. K. M., & Mousis, O. 2012, *ApJL*, **759**, L40
- Mandel, K., & Agol, E. 2002, *ApJL*, **580**, L171
- Mann, A. W., Gaidos, E., Lépine, S., & Hilton, E. J. 2012, *ApJ*, **753**, 90
- Matsumura, S., Peale, S. J., & Rasio, F. A. 2010, *ApJ*, **725**, 1995
- Mayor, M., Marmier, M., Lovis, C., et al. 2011, arXiv:1109.2497
- McQuillan, A., Mazeh, T., & Aigrain, S. 2013, *ApJL*, **775**, L11
- Moorhead, A. V., Ford, E. B., Morehead, R. C., et al. 2011, *ApJS*, **197**, 1
- Morton, T. D., & Johnson, J. A. 2011a, *ApJ*, **729**, 138
- Morton, T. D., & Johnson, J. A. 2011b, *ApJ*, **738**, 170
- Nagasawa, M., & Ida, S. 2011, *ApJ*, **742**, 72
- Naoz, S., Farr, W. M., Lithwick, Y., Rasio, F. A., & Teyssandier, J. 2011, *Natur*, **473**, 187
- Naoz, S., Farr, W. M., & Rasio, F. A. 2012, *ApJL*, **754**, L36
- Nesvorný, D., Kipping, D. M., Buchhave, L. A., et al. 2012, *Sci*, **336**, 1133
- Ofir, A., & Dreizler, S. 2013, *A&A*, **555**, A58
- Ofir, A., Dreizler, S., Zechmeister, M., & Husser, T.-O. 2014, *A&A*, **561**, A103
- Penev, K., Jackson, B., Spada, F., & Thom, N. 2012, *ApJ*, **751**, 96
- Person, D. A. 2012, [http://freepages.misc.rootsweb.ancestry.com/~hallssofjamaica/Kepler\\_long\\_dur.pdf](http://freepages.misc.rootsweb.ancestry.com/~hallssofjamaica/Kepler_long_dur.pdf)
- Petrovich, C. 2014, *ApJ*, in press (arXiv:1405.0280)
- Petrovich, C., Tremaine, S., & Rafikov, R. 2014, *ApJ*, **786**, 101
- Plavchan, P., Bilinski, C., & Currie, T. 2014, *PASP*, **126**, 34
- Quinn, S. N., White, R. J., Latham, D. W., et al. 2012, *ApJL*, **756**, L33
- Rafikov, R. R. 2006, *ApJ*, **648**, 666
- Rasio, F. A., & Ford, E. B. 1996, *Sci*, **274**, 954
- Rogers, T. M., & Lin, D. N. C. 2013, *ApJL*, **769**, L10
- Rogers, T. M., Lin, D. N. C., & Lau, H. H. B. 2012, *ApJL*, **758**, L6
- Santerne, A., Díaz, R. F., Moutou, C., et al. 2012, *A&A*, **545**, A76
- Sari, R., & Goldreich, P. 2004, *ApJL*, **606**, L77
- Schmitt, J. R., Wang, J., Fischer, D. A., et al. 2013, arXiv e-prints
- Slawson, R. W., Prša, A., Welsh, W. F., et al. 2011, *AJ*, **142**, 160
- Socrates, A., & Katz, B. 2012, arXiv:1209.5723
- Socrates, A., Katz, B., & Dong, S. 2012a, arXiv:1209.5724
- Socrates, A., Katz, B., Dong, S., & Tremaine, S. 2012b, *ApJ*, **750**, 106
- Steffen, J. H., Ragozzine, D., Fabrycky, D. C., et al. 2012, *PNAS*, **109**, 7982
- Takeda, G., Ford, E. B., Sills, A., et al. 2007, *ApJS*, **168**, 297
- Tenenbaum, P., Christiansen, J. L., Jenkins, J. M., et al. 2012, *ApJS*, **199**, 24
- Tenenbaum, P., Jenkins, J. M., Seader, S., et al. 2013, *ApJS*, **206**, 5
- Tremmel, M., Fragos, T., Lehmer, B. D., et al. 2013, *ApJ*, **766**, 19
- Tsang, D. 2014, *ApJ*, **782**, 112
- Tsang, D., Turner, N. J., & Cumming, A. 2014, *ApJ*, **782**, 113
- Udry, S., Mayor, M., & Santos, N. C. 2003, *A&A*, **407**, 369
- Wang, J., Fischer, D. A., Barclay, T., et al. 2013, *ApJ*, **776**, 10
- Ward, W. R. 1997, *Icar*, **126**, 261
- Winn, J. N., Fabrycky, D., Albrecht, S., & Johnson, J. A. 2010, *ApJL*, **718**, L145
- Wittenmyer, R. A., O'Toole, S. J., Jones, H. R. A., et al. 2010, *ApJ*, **722**, 1854
- Wright, J. T., Fakhouri, O., Marcy, G. W., et al. 2011, *PASP*, **123**, 412
- Wright, J. T., Marcy, G. W., Howard, A. W., et al. 2012, *ApJ*, **753**, 160
- Wright, J. T., Upadhyay, S., Marcy, G. W., et al. 2009, *ApJ*, **693**, 1084
- Wu, Y., & Lithwick, Y. 2011, *ApJ*, **735**, 109
- Wu, Y., & Murray, N. 2003, *ApJ*, **589**, 605
- Youdin, A. N. 2011, *ApJ*, **742**, 38
- Zahn, J.-P. 1975, *A&A*, **41**, 329